The impact of new experimental data on the global nNNPDF fit

THESIS

submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE
in
THEORETICAL PHYSICS

Author :	Gijs van Weelden
Student ID :	1528408
Supervisor :	Dr. Alexey Boyarsky
2^{nd} corrector :	Dr. Juan Rojo

Leiden, The Netherlands, July 10, 2020

The impact of new experimental data on the global nNNPDF fit

Gijs van Weelden

Instituut-Lorentz, Leiden University
P.O. Box 9500, 2300 RA Leiden, The Netherlands

July 10, 2020

Abstract

Parton distribution functions (PDFs) are vitally important for high energy physics calculations. Vast amounts of experimental evidence have shown that scattering processes involving nuclei cannot be solved using the free-nucleon formalism of perturbative QCD and therefore, a separate empirical determination of the nuclear modification of PDFs is necessary. Because the shape and size of nuclear modification are theoretically unmotivated, the NNPDF collaboration uses a neural network to achieve a model-independent parametrisation. In this thesis, we include new Z boson production data from pPb collisions into the NNPDF framework and examine its impact on the quality of the fit. We will also discuss the phenomenological implications of prompt photon production data in pPb collisions.

Contents

Chapter 1

Introduction

Parton distribution functions (PDFs) are vitally important objects for high energy physics calculations ([1]; [2]), their applications ranging from studying the quark-gluon plasma to the structure of the proton. As non-perturbative objects, there currently does not exist a good determination of PDFs from first principles, necessitating extraction from experimental data ([3]). Furthermore, it has become clear that scattering processes involving nuclei cannot be solved using the free-nucleon formalism of perturbative QCD, necessitating a separate empirical determination of nuclear parton distribution functions (nPDFs).

There are many approaches to determining nPDFs. Some notable recent nPDF determinations are DSSZ ([4]), KA15 ([5]), nCTEQ15 ([6]), EPPS16 ([7]) and TUJU19 ([8]; [9]). This thesis, however, follows the formalism constructed by the NNPDF collaboration ([10]–[51]), where a neural network is used for a model-independent PDF parametrisation.

The first version of the NNPDF approach to nuclear PDFs, nNNPDF1.0, was published in 2019 ([45]). Recently, an improved version of the nPDF determination was published: nNNPDF2.0 ([52]). This version includes many new data sets and displays a much better quark flavour separation than its predecessor. The results of this thesis, as presented in chapter 5, have been incorporated in that paper.

This thesis is concerned with adding two datasets into the NNPDF framework for extracting nPDFs: Z boson production in the CMS detector at $\sqrt{s_{NN}} = 5.12$ TeV and prompt photon production in the ATLAS detector at $\sqrt{s_{NN}} = 8.16$ TeV, both from pPb collisions.

The structure of this thesis is as follows. In chapter 2, we give a short summary of QCD, the theory of the strong interaction, and show how the concept of PDFs arises from a number of QCD processes. Then, we will discuss the concept of nuclear modification and how it leads to a separate nuclear PDF determination. We will also briefly explore the differences between the various nPDF collaborations mentioned above. In chapter 3, we discuss the basic formalism of a deep neural network and how it operates. In chapter 4, we discuss the NNPDF fitting methodology and we present our results in chapter 5. Finally, we give a summary and outlook in chapter 6.

Chapter 2

QCD: a short summary

In the 1950s, an increasingly large number of hadrons was being discovered due to experimental advances. In order to explain the vast amount of observed particles, Murray Gell-Mann (53) and George Zweig (54) proposed that these hadrons were made up of three flavours of quarks: up (u), down (d), and strange (s). There was an initial friction with the spin-statistics theorem, which led to the introduction of the colour quantum number related to an $SU(3)$ symmetry. (55)

In 1973, Kobayashi and Maskawa (56) proposed the existence of three more flavours of quarks: charm (c), bottom (b) and top (t). Also that year, Gross and Wilczek (57), and Politzer (58) discovered that the $SU(3)$ symmetry exhibited both quark binding and asymptotic freedom. This discovery led to the strong interaction being modelled as a theory of quarks with colour charges. The $SU(3)$ quanta are referred to as gluons and the theory as quantum chromodynamics (QCD). Quarks and gluons will be referred to from here on out collectively as "partons".

In this chapter, we will briefly outline the basic formulation of QCD. We will discuss the running coupling and how it leads to colour confinement and asymptotic freedom, and discuss the concept of parton distribution functions and their role in calculating certain observables. For a more detailed description of these subjects, we refer the reader to, e.g., references (55; 59).

A short note on some conventions used in this chapter. Feynman diagrams are drawn with time going from left to right. We will also use the "natural units" convention $c = \hbar = 1$ and the Einstein summation convention with Greek letters indicating four coordinates, e.g. $\mu, \nu = 0, 1, 2, 3$.

2.1 Basic formulation

The Lagrangian of QCD is given by:

$$\mathcal{L}_{QCD} = \bar{q}_i(i\not{D} - m_i)q_i - \frac{1}{4}F^A_{\mu\nu}F^{A\,\mu\nu} \tag{2.1}$$

Here, q_i and $\bar{q}_i$ are the quark and antiquark fields of flavour i with mass m_i, g is the coupling strength and $F^A_{\mu\nu}$ is the gauge field strength tensor of the gluon field A^A_μ. The covariant derivative D_μ is contracted with the Dirac

γ matrices $\not{D} = \gamma^\mu D_\mu$, indicating the fermionic nature of the quarks, and is defined as:

$$(D_\mu)_{ab} = \delta_{ab}\partial_\mu + ig(t^A A^A_\mu)_{ab} \tag{2.2}$$

where a, b are colour indices in the fundamental representation (a, b = r,g,b) and A is a colour index in the adjoint representation: $A = 1, 2, ..., 8$. The matrices t^A are the generators of $SU(3)$ in the fundamental representation and $t^A = \lambda^A/2$ where λ^A are the Gell-Mann matrices, a QCD analogue of the Pauli matrices. The generator matrices t^A obey the commutation relations:

$$[t^A, t^B] = if^{ABC}t^C \tag{2.3}$$

where f^{ABC} are the structure constants of $SU(3)$. We can now write the expression for the field strength tensor as:

$$F^A_{\mu\nu} = \partial_\mu A^A_\nu - \partial_\nu A^A_\mu + gf^{ABC}A^B_\mu A^C_\nu \tag{2.4}$$

The third term in $F^A_{\mu\nu}$ is where QCD's non-abelian character is seen: it gives rise to 3-gluon and 4-gluon vertex interactions. This non-abelian character of QCD gives it one its most important features: the running coupling.

2.2 The running coupling

Using a standard Quantum Field Theoretical approach, one can use the QCD Lagrangian to calculate observables perturbatively, order by order in $\alpha_s = g^2/4\pi$. The β function for an $SU(3)$ symmetry with n_f fermion flavours in the representation is:

$$\beta(g) = \frac{-g^3}{(4\pi)^2}\left(11 - \frac{2}{3}n_f\right) = \frac{-g^3}{(4\pi)^2}\beta_0 \tag{2.5}$$

In terms of α_s, this becomes:

$$\beta(\alpha_s) = \frac{-\alpha_s^2}{2\pi}\beta_0 \tag{2.6}$$

From this, we can derive the running of the coupling:

$$\alpha_s(E) = \frac{1}{\beta_0 \ln \frac{E^2}{\Lambda^2_{QCD}}} \tag{2.7}$$

We see that the coupling α_s is a function of the energy E of the interaction and $\Lambda_{QCD} \sim 200$ MeV, the QCD energy scale. For $E \leq \Lambda_{QCD}$, $\alpha_s \gg 1$, the quarks are tightly bound together, whereas for $E \gg \Lambda_{QCD}$, $\alpha_s \ll 1$ and the quarks become essentially free particles. This last case is referred to as asymptotic freedom and it allows us to apply perturbation theory techniques to QCD in e.g. collider experiments that are performed at high energies. For low energies, the large value of α_s prevents quarks from being observed in isolation. This is referred to as colour confinement.

2.3 Calculating observables

For the calculation of physical observables, we will consider three types of experiments. The first is Deep Inelastic Scattering (DIS), where a lepton scatters off a hadron. The second is the Drell-Yan process (60). The third process we will discuss is prompt photon production.

Let us consider a lepton scattering off a proton: $\ell + p \to \ell' + X$, where X is an unobserved hadron. See figure 2.1 for two leading order diagrams for this process. This process is referred to as deep inelastic scattering if the interaction energy and the mass of the outgoing hadron X are both much greater than the proton mass. There are two separate cases to consider: neutral current (NC) and charged current (CC), characterised by the (electric) charge of the exchanged boson. In the left diagram of figure 2.1, the boson (a photon in this case) is electrically neutral (NC DIS), whereas in the right diagram, the W^+ boson has an electric charge (CC DIS). Also note that from the right diagram, we see that the incoming lepton need not be the same as the outgoing lepton. Similarly, the 'incoming' parton can be a different flavour than the 'outgoing' parton.

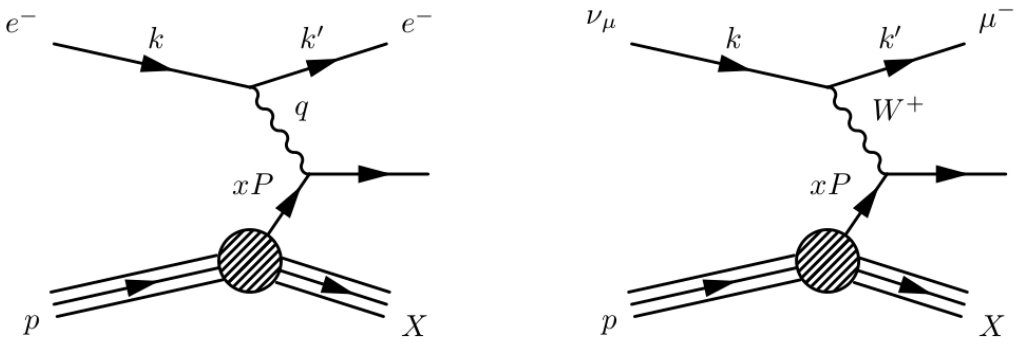

Figure 2.1: Leading order diagrams for a lepton scattering off a hadron via exchange of a virtual boson. The left diagram shows Neutral Current DIS with an electron scattering off a quark via photon exchange. The right shows Charged Current DIS with $\nu_\mu p \to \mu^- X$ via W^+ boson exchange.

For a proton momentum P, we define the momentum of the interacting parton as a fraction x of this momentum. The exchanged energy (i.e. interaction energy) is then $Q^2 = -q^2$ with q the momentum of the exchanged boson. We can now encode the probability for the interacting parton of flavour i (here, a quark) to have momentum fraction x, in a function $f_i(x, Q^2)$, called a parton distribution function (PDF). [1] We use the PDF for the calculation of observables, such as the F_2 structure function:

[1] Strictly speaking, the PDF is not the probability density, but the number density, due to the choice of normalisation.

$$F_2(x,Q^2) = \sum_i^{n_f} C_i(x,Q^2) \otimes f_i(x,Q^2) \tag{2.8}$$

$$= x \sum_i^{n_f} \int_x^1 \frac{dx'}{x'} C_i(x/x',Q^2) f_i(x',Q^2) \tag{2.9}$$

Here, the coefficients C_i are process-dependent functions that can be perturbatively calculated and $\otimes$ is the Mellin convolution, as defined above. This result can be derived using the factorisation theorem (61) and it shows us that the structure function is made up of a perturbative part (C_i) that governs the short distance behaviour and a non-perturbative PDF that governs the long distance behaviour. (3)

NC DIS experiments are only sensitive to one type of quark PDF combination (at leading order). In order to separate the quark flavours, we can use CC DIS measurements, which is sensitive to different types of quark PDF combinations (52). Alternatively, we can consider gauge boson production in hadronic processes. These are processes of the Drell-Yan family of interactions, shown in figure 2.2. The cross-section for such a process is given by:

$$\sigma = \sum_{ij} \int dx_1 dx_2 f_i(x_1,Q^2) f_j(x_2,Q^2) \hat{\sigma}_{ij}(x_1,x_2,Q^2) \tag{2.10}$$

where the two interacting partons of flavour i and j have momentum x_1P_1 and x_2P_2, respectively, and $\hat{\sigma}_{ij}(x_1,x_2,Q^2)$ is the partonic cross-section (3; 62). As with DIS, we see the observable is a function of a perturbative part ($\hat{\sigma}_{ij}$) and non-perturbative PDFs.

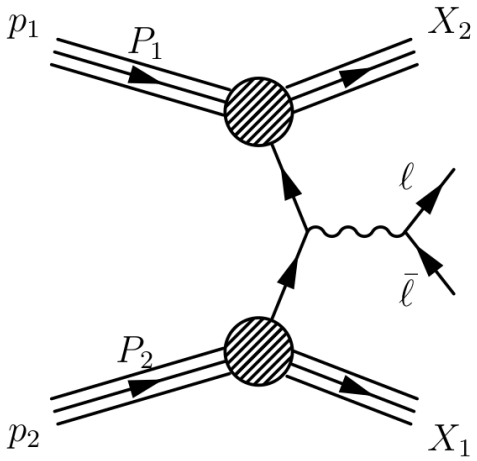

Figure 2.2: Feynman diagram for a Drell-Yan process. A quark-antiquark pair annihilate to produce a lepton-antilepton pair via a gauge boson. Note that the gauge boson can be neutral (γ, Z) or charged ($W^{\pm}$), depending on the flavours of the interacting quarks.

Finally, let us consider prompt photon production. At leading order, prompt photons are produced via QCD Compton scattering or $q\bar{q}$ annihilation, shown in figure 2.3. Photon production is an important QCD process, for

the following reasons. Prompt photons are useful in studying the quark-gluon plasma, as they can traverse it without being modified, due to being QCD neutral. Additionally, as can be seen in figure 2.3, prompt photon production is sensitive to the gluon content of the proton already at leading order, whereas DIS and Drell-Yan processes only sense the gluon beyond leading order. Including photon production processes in our analysis will therefore significantly impact the quality of the gluon PDF fits.

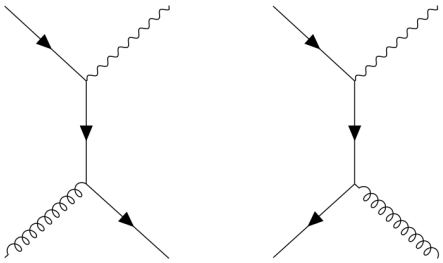

Figure 2.3: Prompt photon production at leading order via QCD Compton scattering (left) and $q\bar{q}$ annihilation (right).

2.4 PDF evolution

In the equations above, we have written the PDFs as a function of Q^2 without further comment. It is interesting to know how the PDFs evolve with Q^2 and, surprisingly, this evolution is governed by perturbative QCD (for $Q^2 \geq 1\,\mathrm{GeV}^2$), despite the PDF's non-perturbative nature ([63]–[65]), and can be derived from the renormalisation group equation. The equations describing the Q^2 evolution of the PDFs are called the DGLAP equations and are given by:

$$\frac{d}{d\log Q} f_g = \frac{\alpha_s(Q^2)}{2\pi} \int_x^1 \frac{dz}{z} \Bigg(P_{g\leftarrow q}(z) \sum_i^{n_f} \left[f_i(x/z, Q^2) + \bar{f}_i(x/z, Q^2) \right] + P_{g\leftarrow g}(z) f_g(x/z, Q^2) \Bigg) \tag{2.11}$$

$$\frac{d}{d\log Q} f_i = \frac{\alpha_s(Q^2)}{2\pi} \int_x^1 \frac{dz}{z} \left(P_{q\leftarrow q}(z) f_i(x/z, Q^2) + P_{q\leftarrow g}(z) f_g(x/z, Q^2) \right) \tag{2.12}$$

$$\frac{d}{d\log Q} \bar{f}_i = \frac{\alpha_s(Q^2)}{2\pi} \int_x^1 \frac{dz}{z} \left(P_{q\leftarrow q}(z) \bar{f}_i(x/z, Q^2) + P_{q\leftarrow g}(z) f_g(x/z, Q^2) \right) \tag{2.13}$$

where the sum over i runs over all n_f quark flavours, f_i is the PDF for flavour i, and $\bar{f}_i$ is the PDF for the respective antiquark. The splitting functions $P_{i\leftarrow j}$ give the probability for a parton of type j to emit a collinear parton of type i with momentum fraction xz. The splitting functions are given by:

$$P_{q\leftarrow q}(z) = \frac{4}{3}\left[\frac{1+z^2}{(1-z)_+} + \frac{3}{2}\delta(1-z)\right] \tag{2.14}$$

$$P_{g\leftarrow q}(z) = \frac{4}{3}\left[\frac{1+(1-z)^2}{z}\right] \tag{2.15}$$

$$P_{q\leftarrow g}(z) = \frac{1}{2}\left[z^2+(1-z)^2\right] \tag{2.16}$$

$$P_{g\leftarrow g}(z) = 6\left[\frac{1-z}{z} + \frac{z}{(1-z)_+} + z(1-z) + \left(\frac{11}{12} - \frac{n_f}{18}\right)\delta(1-z)\right] \tag{2.17}$$

where we $1/(1-z)_+$ is defined as $1/(1-z)$ for $z<1$ and has a singularity at $z=1$ such that:

$$\int_0^1 dz \frac{f(z)}{(1-z)_+} = \int_0^1 dz \frac{f(z)-f(1)}{(1-z)} \tag{2.18}$$

Up until now, we have worked with PDFs of physical quark flavours. This is known as the physical or flavour basis. Because all quarks couple to the gluon, using the DGLAP equations is quite complicated. Therefore, it is convenient to apply a change of basis that decouples the non-singlet combinations of parton distributions from the gluon (66). Defining $q_i^{\pm} = q_i \pm \bar{q}_i$, we can now write:

$$g = g, \quad \Sigma = \sum_i^{n_f} q_i^+, \quad V = \sum_i^{n_f} q_i^-, \quad q_{ij}^{\pm} = q_i^{\pm} - q_j^{\pm} \tag{2.19}$$

where Σ is the (only) quark singlet distribution, V is the valence distribution and $q_{ij}^{\pm}$ are non-singlet distributions. We can now compute the full evolution basis by taking linear combinations of the non-singlet distributions and we will find a basis $\{g, \Sigma, V, V_3, V_8, V_{15}, V_{24}, V_{35}, T_3, T_8, T_{15}, T_{24}, T_{35}\}$. In this work, the g, Σ, V, V_3, T_3 and T_8 PDFs are of importance. So, in addition to the definitions above, we explicitly define V_3, T_3 and T_8:

$$V_3 = u^- - d^- \quad T_3 = u^+ - d^+ \quad T_8 = u^+ + d^+ - 2s^+ \tag{2.20}$$

Now, the DGLAP equations transform for the non-singlet distributions:

$$\frac{d}{d\log Q^2} q_{NS} = \frac{\alpha_s(Q^2)}{2\pi}\int_x^1 \frac{dz}{z} P_{NS}(z) q_{NS}(x/z, Q^2) \tag{2.21}$$

where $P_{NS} = P_{q\leftarrow q}$ at leading order. For the singlet and gluon, the DGLAP equations become:

$$\frac{d}{d\log Q^2}\begin{pmatrix}\Sigma(x,Q^2)\\ g(x,Q^2)\end{pmatrix} = \frac{\alpha_s(Q^2)}{2\pi}\int_x^1 \frac{dz}{z}\begin{pmatrix}P_{q\leftarrow q} & P_{q\leftarrow g}\\ P_{g\leftarrow q} & P_{g\leftarrow g}\end{pmatrix}\begin{pmatrix}\Sigma(x/z,Q^2)\\ g(x/z,Q^2)\end{pmatrix} \tag{2.22}$$

2.5 Nuclear modification

Up until now, we have not made any distinction between partons originating from a free or bound nucleon. In reality, however, these two differ significantly, which is surprising, because the nuclear binding effects are of the order of MeV, while the energy scale of nuclear processes is GeV (67). Vast amounts of experimental evidence have shown that the free-nucleon formalism of perturbative QCD is insufficient to describe the nuclear modification of PDFs (1–3) and while there are a number of theoretical models aiming to explain nuclear modification from first principles, a consensus has yet to be reached (67). Therefore, a separate extraction of nuclear PDFs (nPDFs) from experimental data is necessary.

Nuclear modification was first discovered by the European Muon Collaboration at CERN in 1983 (68) in the form of the EMC effect. Specifically, they observed that the nuclear F_2 structure functions (obtained from DIS experiments) are not the same as the sum of the structure functions of their free nucleon constituents. Instead, they found a pronounced deviation, shown in figure 2.4. The size of this effect has been found to increase with A, but is only feebly affected by the interaction energy Q^2. (67)

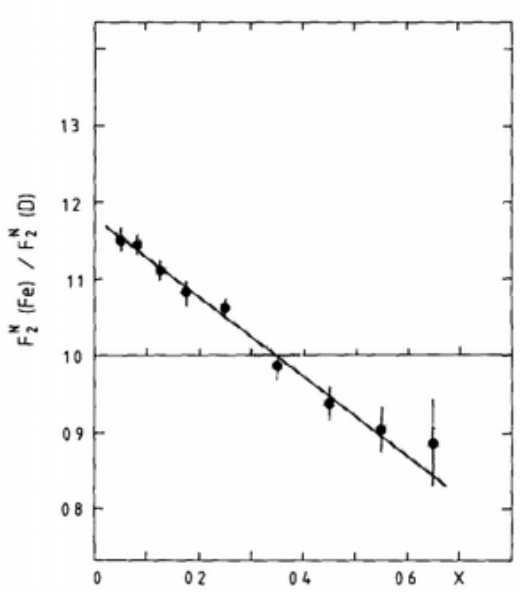

Figure 2.4: Figure from the original EMC paper (68) showing the ratio of the F_2 structure functions for iron (Fe) and deuterium (D). The negative slope of the fitted line is in strong disagreement with the theoretical predictions at the time.

In the following decades, nuclear modification was studied extensively and four distinct regimes (69) of nuclear modification were identified: shadowing ($x \lesssim 0.1$, $R_f^A < 1$), anti-shadowing ($0.1 \lesssim x \lesssim 0.3$, $R_f^A > 1$), EMC effect ($0.3 \lesssim x \lesssim 0.8$, $R_f^A < 1$) and Fermi motion ($x \gtrsim 0.8$, $R_f^A > 1$). We show a schematic representation of these four regimes in figure 2.5 (adapted from reference (45)), where we define the nuclear modification factor R_f^A as the ratio of the PDF in a nucleus to the PDF in a free nucleon:

$$R_f^A = f^{(N/A)}(x, A)/f^{(N)}(x) \tag{2.23}$$

We use the PDF of an average nucleon N in a nucleus with Z protons and atomic mass number A, defined as:

$$f^{(N/A)} = \frac{Z}{A} f^{(p/A)} + \frac{A-Z}{A} f^{(n/A)} \tag{2.24}$$

where $f^{(\{p,n\}/A)}$ are the PDFs for the proton and neutron in a nucleus with atomic mass number A. In the case of isoscalar symmetry ($A = 2Z$), we observe that these average nucleon nPDFs are equivalent ([52]) to the proton nPDFs for all flavours, except the up and down quark, although the relation is still straightforward in those cases. In the evolution basis, the relation is likewise trivial for all flavours but V_3 and T_3, which are related to their proton counterparts by a factor $2Z/A - 1$.

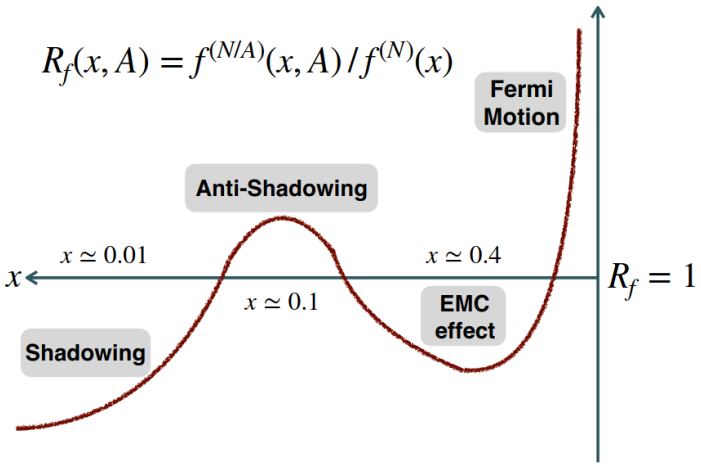

Figure 2.5: Schematic representation of nuclear modification. Indicated are the four distinct regimes of shadowing, anti-shadowing, the EMC effect and Fermi motion. Figure adapted from ([45]).

The first nuclear PDF set was EKS98 ([70]), based on both DIS and Drell-Yan data, quickly followed by the HKM ([71]) set, which included error analysis. Both these sets were at leading order (LO). The first next-to-leading-order (NLO) set was nDS ([72]). With time, the amount of data included in nPDF analyses increased ([3]; [73]) and so did their quality. The most recent nuclear PDF determinations are DSSZ ([4]), KA15 ([5]), nCTEQ15 ([6]), EPPS16 ([7]), TUJU19 ([8]), and nNNPDF2.0 ([52]). Since publication, the nCTEQ15 set has been updated to incorporate $W^{\pm}$ and Z vector boson production from pPb and PbPb collisions ([74]) and the EPPS collaboration has incorporated dijet ([75]) and D-meson ([76]) production in their nPDF set.

We can extract nPDFs by studying processes involving nuclei, where the observable is altered by the nuclear modification of the PDF. As an example, we consider the Drell-Yan cross-section in a pA collision:

$$\frac{d\sigma_{\mathrm{DY}}(y)}{dy} \equiv A\frac{d\sigma_{\mathrm{DY}}^{(N/A)}(y)}{dy} = Z\frac{d\sigma_{\mathrm{DY}}^{(p/A)}(y)}{dy} + (A-Z)\frac{d\sigma_{\mathrm{DY}}^{(n/A)}(y)}{dy} \tag{2.25}$$

where the superscripts $(N/A), (p/A), (n/A)$ indicate that the cross-section corresponds to a collision between a parton from the free proton and a parton from either a bound average nucleon, a bound proton or a bound neutron, respectively. Note that $\sigma^{(p/A)} \neq \sigma^{p}$, the cross-section for a pp collision, but instead is defined by replacing one of the PDFs in equation 2.10 with a nuclear PDF. To illustrate the importance of nuclear PDFs in the modification of observables, we show the effect of nuclear modification (in the EPPS16 nPDF set) on the cross-section of W^- production in pPb collisions (7) in figure 2.6.

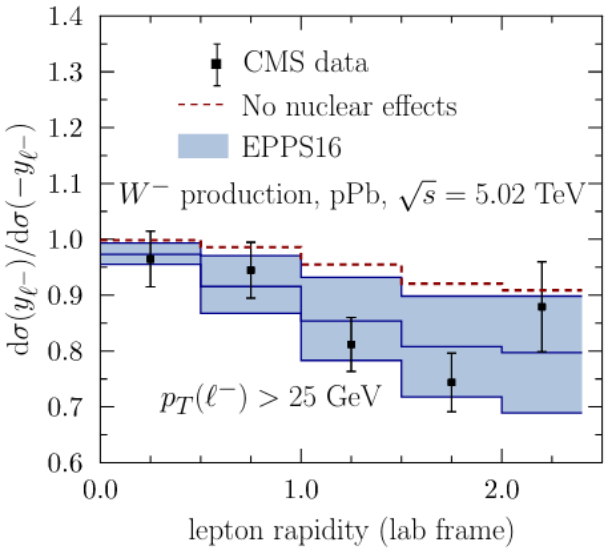

Figure 2.6: Improvement of the theoretical prediction for W^- production in pPb collisions, when including nuclear effects. Figure adapted from (7)

In addition to the modification of observables in pA collisions, we need nPDFs to calculate the initial state of AA collisions. We can also use nPDFs and AA collisions to study the quark-gluon plasma: the hot and dense medium present in the early universe, prior to nucleosynthesis (77). Furthermore, nPDFs are of importance to astroparticle physics in ultra-high energy neutrino scattering processes, probed by neutrino telescopes such as KM3NeT. Lastly, nuclear effects can propagate into the uncertainties of the proton PDF, as many free proton PDF analyses include data on proton-nucleus or lepton-nucleus scattering in their calculations. (3)

2.6 PDF parametrisation

As mentioned before, PDFs are non-perturbative objects. While there have been attempts to derive PDFs from first principles, most notably Lattice QCD, there is currently no reliable approach to do so (3). Therefore, PDFs can be most accurately determined by fitting them from experimental data. In order to do this, we need to establish a parametrisation scheme. It is customary (4–8; 52; 78–80), to parametrise the PDF as follows:

$$f_i(x, Q_0^2) = \mathcal{N} x^{\alpha_i} (1-x)^{\beta_i} \mathcal{I}(x, \{a\}) \tag{2.26}$$

where $\mathcal{N}$ is a normalisation factor that accounts for theoretical constraints, which we will mention later. The factor x^{α_i} governs the low-x behaviour and is derived from Regge Theory (81), whereas the $(1-x)^{\beta_i}$ governs the high-x region and derives from Brodsky-Farrar quark counting rules (82). While some models may predict certain values for the effective exponents α_i and β_i, in practice, they are often fitted from the experimental data (83). The function $\mathcal{I}(x, \{a\})$ is an interpolation function dependent on a set of parameters $\{a\}$. There is no theoretical motivation for the shape of this function and this is where the different (n)PDF collaborations diverge in their methodology and, per extension, their results.

The interpolation function is often chosen to have a polynomial form (8; 78–80) such as $\mathcal{I}(x, \{a\}) = 1 + \gamma\sqrt{x} + \delta x + ...$, with $\{a\} = \{\gamma, \delta, ...\}$, but its shape can be much more complicated (6). Alternatively, the NNPDF collaboration assumes a model-independent approach by parametrising $\mathcal{I}(x, \{a\})$ with a neural network (3; 10; 52), which we will discuss in chapter 4.

Likewise, a parametrisation has to be chosen for the nuclear modification factor R_f^A, which differs between the different nPDF collaborations. One might choose to parametrise it directly (4; 5; 7), similarly to $\mathcal{I}$, or the parameters can be given an A dependent functional form: $\{a(A)\} = \{\gamma(A), \delta(A), ...\}$ (6; 8). Lastly, one can remain model-agnostic and parametrise R_f with a neural network. The nuclear modification factor is then absorbed into the PDF determination of the network. (3; 45; 52)

2.7 Nuclear PDF collaborations

The nPDF landscape is diverse in both the general approach of the problem and the complexity of the models employed. In this section, we will briefly discuss the parametrisation schemes of the aforementioned nPDF collaborations (in chronological order, as per the discussed nPDF sets), with the exception of NNPDF, which we will discuss in chapter 4. The DSSZ, KA and EPPS collaborations do not generate their own free proton PDF sets, but use external sets. The parametrisation employed to construct these free proton PDFs will be discussed separately, in appendix A. Despite the differences in their parametrisation approach, all collaborations use the χ^2 function for their fitting procedure and as the figure of merit for their results. All the nPDF sets discussed in this section, use the Hessian (84) method for uncertainty estimation.

DSSZ: The DSSZ (NLO) nPDF set (4) uses the MSTW08 (79) set as its free proton PDF, which uses a number of different polynomial functions in $\sqrt{x}$ for the various flavours. They parametrise the nuclear modification factor at $Q_0 = 1$ GeV directly as:

$$R_v^A = \varepsilon_1 x^{\alpha_v}(1-x)^{\beta_1}(1+\varepsilon_2(1-x)^{\beta_2})(1+a_v(1-x)^{\beta_3}) \tag{2.27}$$

$$R_s^A = R_v^A \frac{\varepsilon_s}{\varepsilon_1}\frac{1+a_s x^{\alpha_s}}{a_s+1} \tag{2.28}$$

$$R_g^A = R_v^A \frac{\varepsilon_g}{\varepsilon_1}\frac{1+a_g x^{\alpha_g}}{a_g+1} \tag{2.29}$$

where R_v^A is the nuclear modification for the valence quarks and the values for $\varepsilon_1, \varepsilon_2, \varepsilon_s$ and ε_g are fixed by the QCD sum rules (see section 2.8). The A dependence of the other parameters is then parametrised as:

$$\xi = \gamma_\xi + \lambda_\xi A^{\delta_\xi} \tag{2.30}$$

where $\xi \in \{\alpha_v, \alpha_s, \alpha_g, \beta_1, \beta_2, \beta_3, a_v, a_s, a_g\}$. Using the approximation that $\delta_{\alpha_g} = \delta_{\alpha_s}$ and $\delta_{a_g} = \delta_{a_s}$, this leaves a final fit with 25 free parameters.

KA15: The KA15 (NNLO) nPDF set ([5]) uses the JR09 ([80]) free proton PDF set and parametrises the nuclear modification at $Q_0^2 = 2\text{GeV}^2$ directly as:

$$R_i(x, A, Z) = 1 + \left(1 - \frac{1}{A^\alpha}\right) \frac{a_i(A, Z) + b_i(A)x + c_i(A)x^2 + d_i(A)x^3}{(1-x)^{\beta_i}} \tag{2.31}$$

where the A dependence of the parameters for the nuclear modification is parametrised as:

$$a_{\bar{q}}(A) = a_1 A^{a_2} \tag{2.32}$$
$$b_i(A) = b_1 A^{b_2} \tag{2.33}$$
$$c_i(A) = c_1 A^{c_2} \tag{2.34}$$
$$d_i(A) = d_1 A^{d_2} \tag{2.35}$$

with i indices left implicit. KA15 chooses fixed values for a number of parameters. They set $\alpha = 1/3$, due to constraints imposed by nuclear volume and surface contributions and $\beta_v = 0.4, \beta_{\bar{q}} = 0.1, \beta_g = 0.1$, due to a lack of data preventing them from determining these values from the fit. The values for the a_v and a_g parameters are fixed by the QCD flavour and momentum sum rules, respectively. The other parameters are determined via the fitting procedure, yielding a total of 16 free parameters.

nCTEQ15: The nCTEQ15 (NLO) set ([6]) opts for an exponential interpolation function while simultaneously fitting the ratio of $\bar{u}$ and $\bar{d}$ quarks:

$$x f_i^{p/A}(x, Q_0) = c_0 x^{c_1} (1-x)^{c_2} e^{c_3 x} (1 + e^{c_4} x)^{c_5} \tag{2.36}$$
$$\frac{\bar{d}(x, Q_0)}{\bar{u}(x, Q_0)} = c_0 x^{c_1} (1-x)^{c_2} + (1 + c_3 x)(1-x)^{c_4} \tag{2.37}$$

where $i \in \{u_v, d_v, g, \bar{u} + \bar{d}, s + \bar{s}, s - \bar{s}\}$. The nuclear modification is then parametrised at $Q_0 = 1.3$ GeV by introducing an A dependence in the fitting parameters c_k:

$$c_k \to c_k(A) \equiv c_{k,0} + c_{k,1}(1 - A^{c_{k,2}}) \quad k = 1, 2, ..., 5 \tag{2.38}$$

In total, nCTEQ15 allows for ~ 10 free parameters per parton flavour. Due to data limitations, however, they constrain themselves to a fit with 16 free parameters: 7 for the gluon, 4 for the valence u quark, 3 for the valence d quark and 2 for the $\bar{d} + \bar{u}$ quark.

EPPS16: The EPPS16 (NLO) nPDF set ([7]) uses the CT14 ([78]) free proton PDF, which employs a fourth-order polynomial in $y = \sqrt{x}$. However, in order to decorrelate the parameters of this polynomial, they instead fit a linear combination of Bernstein polynomials and translate the fitted parameters back to those of the interpolation function. EPPS16 opts for a direct parametrisation of R_f^A at $Q_0 = m_c = 1.3$ GeV with polynomial functions:

$$R_f^A(x, Q_0^2) = \begin{cases} a_0 + a_1(x - x_a)^2 & x \leq x_a \\ b_0 + b_1 x^\alpha + b_2 x^{2\alpha} + b_3 x^{3\alpha} & x_a \leq x \leq x_e \\ c_0 + (c_1 - c_2 x)(1 - x)^\beta & x_e \leq x \leq 1 \end{cases} \tag{2.39}$$

where $\alpha = 10 x_a$, x_a is the position of the anti-shadowing maximum, x_e is the position of the EMC minimum and the coefficients a_i, b_i, c_i are determined by the asymptotic small-x limit of R_f^A. Using $y_i = R_f^A(x_i, Q_0^2)$ for $x_i = 0, x_a, x_e$, the A dependence of y_i is parametrised as:

$$y_i = y_i(A_{ref}) \left(\frac{A}{A_{ref}} \right)^{\gamma_i [y_i(A_{ref}) - 1]} \tag{2.40}$$

where $\gamma_i \geq 0$ and $A_{ref} = 12$. The nuclear modification, deviation from $R_f^A = 1$, is now greater for high A, by construction. Lastly, continuity and vanishing first derivatives are required for R_f^A at x_i. In total, the EPPS16 fit has 56 parameters, 36 of which are fixed, leaving 20 free fitting parameters.

TUJU19: Lastly, let us consider the TUJU19 (NLO and NNLO) nPDF set ([8; 9]) which uses a simple second order polynomial for the interpolation function at $Q_0^2 = 1.69\,\text{GeV}^2$:

$$x f_i^{p/A} = c_0 x^{c_1} (1 - x)^{c_2} (1 + c_3 x + c_4 x^2) \tag{2.41}$$

As in the nCTEQ15 analysis, the nuclear modification is parametrised by introducing an A dependence into the fitting parameters:

$$c_k \to c_k(A) \equiv c_{k,0} + c_{k,1}(1 - A^{c_{k,2}}) \tag{2.42}$$

where $c_{k,0}$ is kept fixed for all flavours based on the free proton fit, and the nuclear parameters $c_{k,1}, c_{k,2}$ are fitted for each flavour. Under the TUJU19 fitting assumptions, this equates to a fit with 16 free nuclear parameters in total.

2.8 Theoretical constraints

As mentioned above, PDFs are the probability distributions that dictate the momentum of partons as a fraction x of the hadron momentum. Although this interpretation is no longer valid when we move beyond leading order ([85]), the PDF is still constrained ([86]) by a normalisation constraint, the momentum sum rule, given in equation 2.43, due to conservation of energy. Additionally, baryon number conservation yields a flavour or valence sum rule, given in equation 2.44.

$$\sum_i \int_0^1 dx x f_i(x, Q_0^2, A) = 1 \tag{2.43}$$

$$\int_0^1 dx \left(f_i(x, Q_0^2, A) - \bar{f}_i(x, Q_0^2, A) \right) = n_i \tag{2.44}$$

with n_i the number of quarks of flavour i. Note that these sum rules are valid for all values of A and need only be computed for a single energy scale Q_0, as the DGLAP equations guarantee their validity at all $Q > Q_0$ (45). Switching from the physical basis to the evolution basis, the momentum and valence sum rules become:

$$\int_0^1 dx x(\Sigma + g) = 1 \tag{2.45}$$

$$\int_0^1 dx V = \sum_i n_i \tag{2.46}$$

Although the validity of the sum rules has been called into question for the nuclear case (87), no definitive evidence for this has been found. The NNPDF collaboration has recently found their nPDF fits to satisfy the sum rules (within uncertainties) (52) even if they were not imposed. In addition to the sum rules mentioned above, there are some theoretical constraints on the allowed sizes and shapes of PDFs.

For $x \to 1$, any PDF must go to zero (83). If a parton were to possess all of the momentum of a proton or neutron, it would be a free particle, which is forbidden by colour confinement.

While PDFs can, in general, be negative, hadronic observables are positive definite (3). One can ensure this positivity constraint in a number of ways. One can choose the parametrisation such that positivity is guaranteed or simply discard the PDF parameter configurations that lead to negative observables.

All nuclear PDFs are constrained for $A = 1$ by the proton PDF. Again, this can be constrained by the choice of parametrisation of the nPDF (3). Alternatively, one can fit the $A = 1$ PDF alongside the other nuclei, compare it to a proton PDF prior and discard the fits that do not agree within its uncertainties (45). The latter approach results in smaller uncertainties for nuclei with low A.

Chapter 3

Neural Networks

The concept of artificial neural networks was first coined by McCulloch and Pitts in 1943 (88). The idea was to mimic human intelligence by using a structure of connected neurons. Neural networks are well suited for non-linear regression and classification problems, even in cases where other machine learning techniques break down. Despite this vast potential, neural networks fell out of favour due to their unfeasably high computational cost (89). However, due to advances in computation in recent decades, the potential of neural networks is now accessible and they are being used for many different purposes, ranging from natural language processing to theoretical physics problems. (90)

In this chapter, we will discuss some of the basic properties of neural networks: their structure and how they learn. For an in-depth look at (deep) neural networks, we refer the reader to reference (91; 92).

Generally, we distinguish three different types of learning for a neural network: supervised, unsupervised and reinforcement learning. In supervised learning, the data the network is trained on is labelled, i.e., the desired outcome is known. Supervised learning is used in e.g. classification or regression problems (93). Unsupervised learning has the network find correlations within the data without any preconstructed labels. This form of learning is a powerful data compression or clustering tool (94). Finally, reinforcement learning teaches a network to interact with its environment. A prime example would be a network learning to play a game (95). While the considerations below are quite general, we focus in this work on supervised learning, which is the type of learning employed in the NNPDF framework.

3.1 Network Architecture

A neural network consists of layers made up of individual neurons (elements) which are connected to the neurons in adjacent layers with individual weights, see figure 3.1. The weights parametrise the sensitivity of a neuron to the values of each of the neurons in the previous layer. The network has an input layer, an output layer, and can have an arbitrary number of intermediate, hidden, layers in-between. Each neuron has an individual bias, which parametrises its sensitivity to the total input it receives from the previous layer. Each layer (apart from the input layer) has an activation function that introduces the non-linear behaviour.

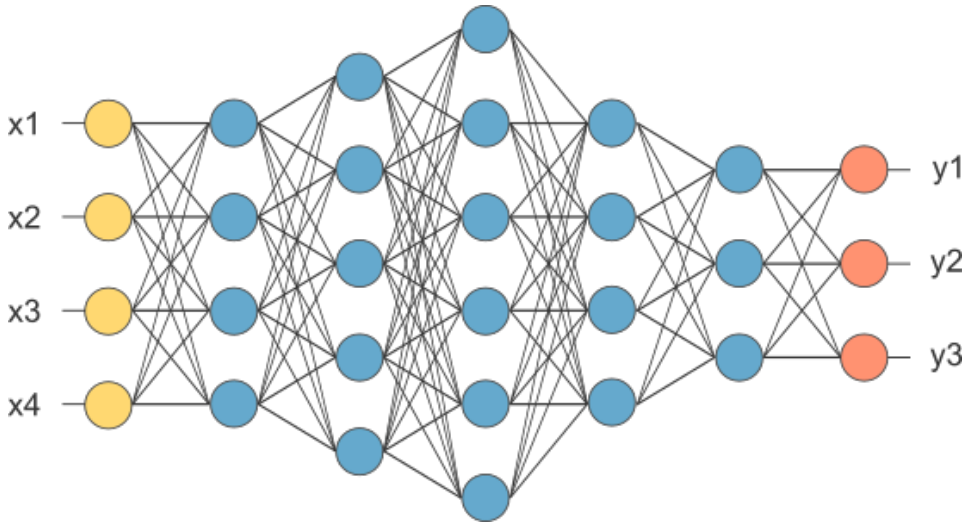

Figure 3.1: A neural network made up of an input layer (yellow) with 4 neurons, five hidden layers (blue) of various sizes and an output layer (red) with 3 neurons. The lines connecting the neurons signify the weights. Figure adapted from [96]

The value z_i^ℓ of neuron i in layer ℓ is given by:

$$z_i^\ell = a^\ell \left(\sum_j w_{ij}^{\ell-1} z_j^{\ell-1} - b_i^\ell \right) \tag{3.1}$$

where $w_{ij}^{\ell-1}$ is the weight connecting neuron j in layer $\ell - 1$ to neuron i in layer ℓ, $a^\ell(z)$ is the activation function in layer ℓ, and b_i^ℓ is the bias of neuron i in layer ℓ (sometimes referred to as a threshold). Using this update rule, the values of the input vector are propagated to the end of the network into the output vector $\mathbf{z}^L$. We can then relate our output vector to the desired output vector $\mathbf{y}$ and define a cost or loss function $C(\mathbf{y}, \mathbf{z}^L)$, in such a way that an optimal result coincides with a minimum value of the cost function.

Before any calculation can be performed, we need to initialise the weights and biases of the network. The weights are commonly initialised randomly, although the used probability distribution may vary [91], whereas the biases are initialised at 0, as initialising the biases at a non-zero value may lead to much longer run times.

3.2 Learning

After having calculated the output of the network, the value of the cost function can be determined. Now, we want to slightly alter our network, in order to achieve a better result next run. For this, we use a stochastic gradient descent algorithm [91]. After each step, we change the weights and biases such that we move down along the gradient of the cost function, i.e., to a better result. We then update the weights and biases by:

$$\delta W_{ij}^{\ell} = -\eta \frac{\partial C}{\partial W_{ij}^{\ell}} \tag{3.2}$$

$$\delta b_{i}^{\ell} = -\eta \frac{\partial C}{\partial b_{i}^{\ell}} \tag{3.3}$$

where η is the learning rate, a parameter that governs the size of the steps taken along the descent. The learning rate can be a constant, but is often allowed to vary over time. The advantage of this is clear: while a high learning rate might be advantageous at the start of learning, as a minimum is approached, a high learning rate will cause the network to overshoot, thus delaying the network or even outright preventing it from reaching a minimum.

Because the network aims to minimise the cost function with every iteration (or epoch), there is a risk of overfitting: the network fitting to the noise of the data instead of the underlying distribution. While the cost function does not inherently contain any information on whether the network is overfitting, we can use it as a measure of overfitting by making use of a validation set (91). Instead of training our network on all of the available data, we split the data and train our network on part of it: the training set. Then, after each epoch, we use the trained network to fit the data in the validation set and record the value of the cost function for that set. Now, the objective becomes not to minimise the cost function on the training set, but on the validation set. As illustrated in figure 3.2, the error on the training data continues to decrease (higher accuracy), but the error of the validation set (representative of data the network has never seen before), has passed its minimum value (maximum accuracy), signifying overfitting.

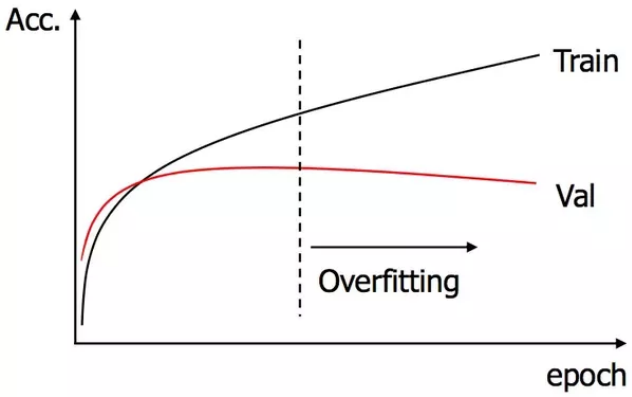

Figure 3.2: Accuracy (inverse error) of the training and validation set. The vertical dashed line indicates the optimal stopping point for the network training: the accuracy of the network on the validation set is maximum. Figure adapted from (97)

3.3 Optimisers

There are various ways to improve the gradient descent algorithm. One popular optimisation is the addition of momentum (91; 98–100). Analogous to the

physical concept, the learning rule is updated such that the network moves faster in directions with persistent downward gradients. The learning rule then becomes:

$$v_t = \gamma v_{t-1} - \eta \nabla_\theta C(\theta_t) \tag{3.4}$$
$$\theta_{t+1} = \theta_t - v_t \tag{3.5}$$

where we have introduced the momentum parameter γ, with $0 \leq \gamma \leq 1$, and we have combined all parameters (weights and biases) into the θ_t parameter. One particular form of momentum in gradient descent is Nesterov Accelerated Gradient Descent (NAG) ([99]). In NAG, rather than calculating the gradient at the current parameters, we calculate the gradient at the expected position:

$$v_t = \gamma v_{t-1} - \eta \nabla_\theta C(\theta_t + \gamma v_{t-1}) \tag{3.6}$$
$$\theta_{t+1} = \theta_t - v_t \tag{3.7}$$

Nesterov momentum allows for a larger learning rate η, for the same value of γ, thus allowing for faster convergence ([91]).

Instead of scaling up persistent gradients, we can make the optimisation algorithm more sensitive to sparse parameter regions, by tuning the learning rate to the parameters. This is known as an adaptive gradient or AdaGrad ([101]). Writing the gradient at time t as g_t, the (component-wise) update rules for AdaGrad are:

$$g_{t,i} = \nabla_\theta C(\theta_{t,i}) \tag{3.8}$$
$$\theta_{t+1,i} = \theta_{t,i} - \frac{\eta}{\sqrt{G_{t,ii} + \varepsilon}} g_{t,i} \tag{3.9}$$

where $G_{t,ii}$ are the sums of the squares of the gradients with respect to the parameter θ_i, up to time t, and ε is a small constant to prevent divergences. A problem with the AdaGrad optimiser is that its learning rate rapidly decreases with time. In order to combat this, Hinton ([102]) proposed an alternative optimiser that still uses past gradient information, but is only sensitive to that information for a limited period of time: RMSProp. Instead of the weighted average gradient, RMSProp uses the second moment $s_t = \langle g_t^2 \rangle$. The (component-wise) RMSProp update rule is:

$$s_t = \beta s_{t-1} + (1 - \beta) g_t^2 \tag{3.10}$$
$$\theta_{t+1} = \theta_t - \eta g_t / \sqrt{s_t + \varepsilon} \tag{3.11}$$

where the decay rate β is a constant that governs the averaging time of the gradients. ([91])

The Adaptive Momentum Estimation (Adam) optimiser ([103]) combines the advantages of both AdaGrad and RMSProp by keeping track of both the first and second moment of the gradient. Adam updates the first and second moments as:

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t \tag{3.12}$$
$$s_t = \beta_2 s_{t-1} + (1 - \beta_2) g_t^2 \tag{3.13}$$

where m_t is the first moment with decay rate β_1 and s_t is the second moment with decay rate β_2. Accounting for the fact that we are estimating these moments with a running average, Adam performs a bias correction:

$$\hat{m}_t = \frac{m_t}{1 - (\beta_1)^t} \tag{3.14}$$
$$\hat{s}_t = \frac{s_t}{1 - (\beta_2)^t} \tag{3.15}$$

Now, we can use these bias-corrected moments to update the network parameters as:

$$\theta_{t+1} = \theta_t - \eta \frac{\hat{m}_t}{\sqrt{\hat{s}_t} + \varepsilon} \tag{3.16}$$

Chapter 4

Methodology

In this chapter, we will discuss the NNPDF methodology. We will detail the most important software packages used and the PDF parametrisation using a neural network. We also discuss the initialisation and training of the network.

4.1 Monte Carlo simulation

`MCFM` is a parton-level Monte Carlo (MC) programme for femtobarn processes ([104–106]). This programme simulates partonic processes to give a (differential) cross-section for various processes occurring in hadron-hadron collisions. The version used for this work is `MCFM v6.8`, which can calculate various processes to NLO precision. A full list of the available processes can be found in the `MCFM` documentation ([107]).

`MCFM` allows the user to choose a number of settings to fit the experiment. Most notably, for this work, these include the centre of mass energy of the experiment, and kinematic cuts on quantities such as the pseudorapidity or the transverse momentum.

4.2 `APPLgrids`

Normally, the MC run needs to be repeated for each new input PDF set, which is very computationally expensive. The `APPLgrid` formalism is a solution to this by allowing for the a posteriori inclusion of PDFs into the MC run ([108]). Instead of having the MC run calculate a histogram of cross-sections, it calculates a lookup table of weights in (x, Q^2) that the PDF can subsequently be combined with. This way, the MC calculation needs to be performed only once and can be used in conjunction with any amount of different PDF sets.

4.3 `FK` tables

In order to compute observables with PDFs, one needs to perform complicated convolutions, as shown in equation 2.8 or 2.10 while determining observables. We can simplify this calculation greatly by using `FastKernel` (`FK`) tables: pre-calculated lookup tables that contain all perturbative information (calculated using the constructed `APPLgrid`) and a suitable interpolation basis ([20; 28]).

To illustrate this, let us look at the expression for the F_2 structure function in DIS.

First, let us assume we can write the PDF with two interpolating functions $I_\alpha(x)$ and $I_\beta(Q^2)$, such that:

$$f_i(x, Q^2, A) = \sum_\alpha \sum_\beta f_i(x_\alpha, Q^2_\beta, A) I_\alpha(x) I_\beta(Q^2) \tag{4.1}$$

We can express $f_i(x_\alpha, Q^2_\beta, A)$ at the input energy scale using the interpolated DGLAP operators:

$$f_i(x_\alpha, Q^2_\beta, A) = \sum_j \sum_\gamma \Gamma_{ij,\alpha\beta\gamma} f_j(x_\gamma, Q^2_0, A) \tag{4.2}$$

Now, we can rewrite the F_2 structure function and define the `FK` tables accordingly:

$$F_2(x, Q^2, A) = \sum_i^{n_f} C_i(x, Q^2) \otimes f_i(x, Q^2, A) \tag{4.3}$$

$$= \sum_i^{n_f} C_i(x, Q^2) \otimes \sum_j \sum_{\alpha,\beta,\gamma} \Gamma_{ij,\alpha\beta\gamma} f_j(x_\alpha, Q^2_0, A) I_\beta(x) I_\gamma(Q^2) \tag{4.4}$$

$$= \sum_i^{n_f} \sum_\alpha^{n_x} \texttt{FK}_{i,\alpha}(x, x_\alpha, Q^2_0, Q^2) f_i(x_\alpha, Q^2_0, A) \tag{4.5}$$

Thus, by using `FK` tables, we can replace the convolutions by matrix multiplication, greatly speeding up the computation. In addition, we circumvent the calculation of the complicated integro-differential DGLAP equations 2.21, 2.22 by including them in the `FK` table. For a complete treatment of the FastKernel method, we refer the reader to references (20) and (28).

4.4 Pre-processing data with `buildmaster`

Before we can use our network to fit the data, we have to convert the data into a format that our code is equipped to handle. During this translation of data formats, we must also ensure the uncertainties of the data are propagated correctly into the new format. We use a program called `buildmaster` for this conversion.

The experimental data is conventionally stored in the HEPData (109) database, where it can be downloaded along with the corresponding uncertainty information. For each data set, we then have to construct a filter, which will read the information from the data files and store it in (`C++`) arrays, where it can be then converted to the new data format by the `buildmaster` code.

The filter also determines the treatment of the uncertainties. If necessary, it symmetrises the uncertainties and shifts the data values accordingly. We label the uncertainties depending on whether they are correlated with the experimental data and calculate both their additive and multiplicative form.

Statistical uncertainties are always additive, by their random nature, and the luminosity uncertainty is always multiplicative. For other uncertainties, we must determine from the original publication of the experimental results whether the uncertainties are to be treated as additive or multiplicative. In case this is not clear, the NNPDF policy is to assume the uncertainties are multiplicative for collider experiments, so as to avoid the d'Agostini bias. (19)

4.5 Neural network

The neural network used to fit the nPDFs (for the nNNPDF2.0 nPDF set) has a 3-25-6 architecture with a sigmoid activation function in the hidden layer and a linear activation function in the output layer (52). The three input neurons correspond to x, $\ln 1/x$ and A, and the output neurons correspond to the nPDFs of interest in the evolution basis at the initial energy scale Q_0^2. We use both x and $\ln 1/x$ as input because only using x as input can make the network lose its sensitivity to small x, as the neuron will feed forward a near-zero value for any reasonably sized weight. At low values of x, $\ln 1/x$ can still be ~ 1 and so the network retains its accuracy for small values of x.

It has been shown (45) that PDF fits are stable with respect to this network architecture. This means that if one were to increase the number of neurons in the hidden layer, the fit will change only within statistical fluctuation. This implies the network is sufficiently redundant in its 3-25-6 shape, equating to 256 free parameters (weights and biases). Additionally, the same has been shown to hold for a network with a similar amount of parameters, but with two hidden layers.

NNPDF assumes three active quarks, vanishing strangeness asymmetry, and c and b quarks generated via perturbative evolution. These assumptions imply (52) a six-parton fitting basis $\{u, \bar{u}, d, \bar{d}, s, g\}$ where $s = \bar{s}$. The corresponding evolution basis is then $\{\Sigma, g, V, T_8, V_3, T_3\}$, which is related to the flavour basis via equations 2.19 and 2.20. We parametrise these nPDFs at energy $Q_0 = 1$ GeV as:

$$\begin{aligned}
x\Sigma^{(p/A)}(x,Q_0) &= x^{\alpha_\Sigma}(1-x)^{\beta_\Sigma}\mathrm{NN}_\Sigma(x,A)\\
xg^{(p/A)}(x,Q_0) &= B_g x^{\alpha_g}(1-x)^{\beta_g}\mathrm{NN}_g(x,A)\\
xV^{(p/A)}(x,Q_0) &= B_V x^{\alpha_V}(1-x)^{\beta_V}\mathrm{NN}_V(x,A)\\
xT_8^{(p/A)}(x,Q_0) &= x^{\alpha_{T_8}}(1-x)^{\beta_{T_8}}\mathrm{NN}_{T_8}(x,A)\\
xT_3^{(p/A)}(x,Q_0) &= x^{\alpha_{T_3}}(1-x)^{\beta_{T_3}}\mathrm{NN}_{T_3}(x,A)\\
xV_3^{(p/A)}(x,Q_0) &= B_{V_3} x^{\alpha_{V_3}}(1-x)^{\beta_{V_3}}\mathrm{NN}_{V_3}(x,A)
\end{aligned} \tag{4.6}$$

where NN_i are the output neurons of the network. Note that we fit the bound proton PDFs $f^{(p/A)}$, instead of the average nucleon nPDF $f^{(N/A)}$. The reasons for this are threefold: a straight-forward connection to the $A = 1$ (free proton) boundary condition, avoiding Z dependence of the PDFs, and avoiding Z/A dependence in the sum rules for non-isoscalar nuclei (52). The normalisation factors are determined by the sum rules (equations 2.45 and 2.46) as:

$$B_g(A) = \frac{1 - \int_0^1 dx\, x\Sigma^{(p/A)}(x, Q_0)}{\int_0^1 dx\, xg^{(p/A)}(x, Q_0)} \tag{4.7}$$

$$B_V(A) = \frac{3}{\int_0^1 dx\, V^{(p/A)}(x, Q_0, A)} \tag{4.8}$$

$$B_{V_3}(A) = \frac{1}{\int_0^1 dx\, V_3^{(p/A)}(x, Q_0, A)} \tag{4.9}$$

where the denominators are calculated using equation 4.6 with $B_g = B_V = B_{V_3} = 1$.

4.6 Network initialisation

The weights of the network are initialised via Xavier initialisation (103), which samples from a normal distribution with zero mean and variance of $1/N$, with N the amount of neurons in the previous layer. In addition, the initial values for the input weights for the hidden layer are constrained to be within two standard deviations, which leads to more efficient training (45). The biases are initialised at zero. The effective exponents α_i, β_i are sampled uniformly from the intervals listed below and fitted simultaneously with the other parameters of the network. The values of the effective exponents are always constrained to be within the intervals in brackets. For more information on the treatment of α_i, β_i, we refer the reader to references (45) and (52).

$$\begin{aligned} \alpha_{\{\Sigma, g, T_8, T_3\}} &\in [-1, 1] \quad ([-1, 5]) \\ \alpha_{\{V, V_3\}} &\in [1, 2] \quad ([0, 5]) \\ \beta_{\{\Sigma, g, T_8, T_3, V, V_3\}} &\in [1, 5] \quad ([1, 10]) \end{aligned}$$

We train the network with the χ^2 of the fit as the cost function, combined with additive terms representing the bound proton and positivity boundary conditions, and we use the Adam (110) optimiser, discussed in section 3.3, to perform the stochastic gradient descent. We use standard values for most of the Adam parameters (45): the initial learning rate $\eta = 0.001$, the decay rate of the second moment of past gradients $\beta_2 = 0.999$ and the smoothing parameter $\varepsilon = 10^{-8}$. The only deviation from standard values is the decay rate of the first moment of past gradients $\beta_1 = 0.99$, which is slightly larger than its standard value of 0.9, because this was found to result in better overall performance (45).

The full cost function C is given by the χ^2 of the fit (4.10), the proton boundary condition (4.11) and the positivity penalty for the hadronic observables (4.12):

$$C = \chi^2 \tag{4.10}$$

$$+ \lambda_{BC} \sum_f \sum_i^{N_x} \left(q_f^{(p/A)}(x, Q_0^2, A = 1) - q_f^p(x, Q_0^2) \right)^2 \tag{4.11}$$

$$+ \sum_k^{N_{pos}} \lambda_{pos}^k \sum_j^{N_A} \sum_{i_k}^{N_{dat}^k} \max\left(-\mathcal{F}_{i_k}^k(A_j), 0\right) \tag{4.12}$$

where λ_{BC} and λ_{pos}^k are Lagrange multipliers indicating the weight of each of these conditions. The sum over f runs over all active partons in the evolution basis. The sum over i runs over $N_x = 60$ points with 10 points spread logarithmically between $x = 10^{-3}$ and $x = 0.1$ and 50 points spread linearly from $x = 0.1$ and $x = 0.7$. The proton baseline q_f^p is taken to be a variant of the NNPDF3.1 NLO free proton fit, that excludes heavy nuclear target data. In equation 4.12, we sum over N_{pos} observables $\mathcal{F}_{i_k}^k$, each with N_{dat} data points, for all N_A available values of A_j ([52]). λ_{BC} is set to 10^4 to ensure that the contribution of 4.11 is of the same order as the χ^2, while λ_{pos}^k is manually tuned by observing the optimisation process.

4.7 Central value and uncertainties

In order to improve the quality of the fit, we make use of Monte Carlo generated pseudo-data. We use a MC method to generate so-called replicas of the data to which we can fit a network. Each replica then yields a distinct (n)PDF fit. We determine our central value by taking the median of all these fits and the uncertainties are determined by the distance of our fits to this central value ([19]). Note that we will need to alter our convergence criterion when fitting replicas. For the true data, a $\chi^2/N_{dat} \sim 1$ indicates a good fit: the variance of the fit is of the order of the variance of the data. Because independent errors add in quadrature, the variance of the pseudo-data will be double that of the true data (provided we set the variance of the MC sampling distribution equal to the variance of the data). A good fit should then yield $\chi_k^2/N_{dat} \sim 2$, where the subscript k indicates a single replica. The average fit, however, should again have $\chi^2/N_{dat} \sim 1$. ([2])

Chapter 5

Results

In this thesis, we have incorporated the data of two LHC-based experiments in the NNPDF framework and examined their impact on the overall quality of the global nPDF fits. We have studied Z boson production in pPb collisions at centre of mass energy $\sqrt{s_{NN}} = 5.02$ TeV in the CMS detector (111) and prompt photon production in pPb collisions in the ATLAS detector at $\sqrt{s_{NN}} = 8.16$ TeV (112).

5.1 CMS Z production

This CMS experiment (111) studies the process of a pPb collision producing a Z boson, decaying to a lepton-antilepton pair: $Z \to \ell\bar{\ell}$. This process is a member of the Drell-Yan family discussed in section 2.3. The experiment is performed at a centre of mass energy per nucleon of $\sqrt{s_{NN}} = 5.02$ TeV at an integrated luminosity of $\mathcal{L} = 34.6 \pm 1.2\,\text{nb}^{-1}$. The lepton pseudorapidity is limited to $|\eta^{\ell}_{lab}| < 2.4$ in the lab frame and the lepton minimum transverse momentum is $p^{\ell}_{T} > 20$ GeV. The cross-sections are given as a function of the centre-of-mass rapidity y_{CM}, which is limited to the interval: $-2.8 < y_{CM} < 2.0$.

The first step in studying this experiment, is to use the kinematic cuts listed above to perform a MC simulation with `MCFM`. The full settings used can be found in appendix B. This constructs an `APPLgrid`, which we can convolute with various PDF sets and consequently compare the predicted cross-sections with those given in the reference (experimental) paper. In figure 5.1, we show the comparison between our `APPLgrid` implementation and the values given in the reference, for the CT10nlo (113), EPS09 (114) and DSSZ [1] (4) input PDF sets. We also show the values of the total (integrated) cross-sections, which show agreement within 1%. In the lower plot, we show the ratio of the reference and prediction values normalised w.r.t. the CT10nlo PDF set. This good agreement between our predictions and the reference values validates our `APPLgrid` implementation as a good representation of the experimental data.

Using this `APPLgrid`, we can then generate the `FK` tables and implement the data in the `buildmaster`. On HEPData, we find the (symmetric) total systematic uncertainties, presented as additive, uncorrelated errors, and the statistical uncertainty. Lastly, we also have a luminosity uncertainty of 3.5%.

[1] This set was converted to a LHAPDF (85) set (both Hessian and MC versions) by Emanuele Nocera. We used the MC version. (44; 115)

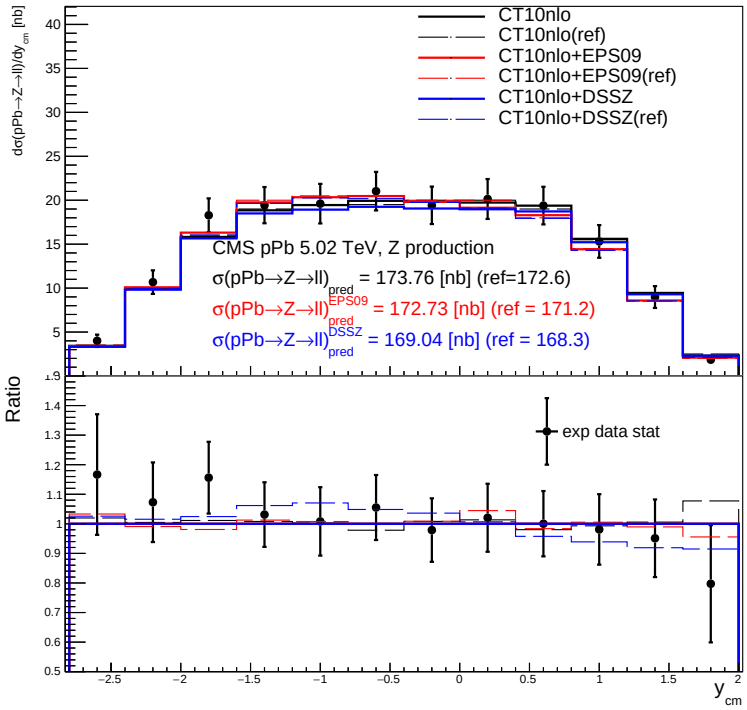

Figure 5.1: Differential cross-section for Z production in pPb collisions at $\sqrt{s_{NN}} = 5.12$ TeV. The solid lines correspond to the predictions calculated by convoluting PDF sets with our `APPLgrid` and the dashed lines are the predictions given in the CMS paper. ([111]) The lower plot shows the ratio of these values normalised w.r.t. the CT10nlo ([113]) predictions.

With our `buildmaster` implementation now complete, we can train the neural network and extract the nPDFs.

The impact of this single data set on the global nNNPDF2.0 fit will be small, as it accounts for $< 1\%$ of the total data points and 12.8% of the total Drell-Yan data. In addition, the effect of CMS Z data will be similar to that of the other Drell-Yan type data. Therefore, it is more instructive to examine the effect of the Drell-Yan data as a whole.

In figure 5.2, we show the nuclear modification factor for lead nuclei for all fitted parton flavours at $Q^2 = 100\,\mathrm{GeV}^2$. We compare the DIS only fit (orange line) to the global nNNPDF2.0 fit (blue line), where the global fit contains both DIS and Drell-Yan data, and the shaded bands correspond to the 90% confidence level. For a full list of the data sets included in the global nNNPDF2.0 fit and their corresponding χ^2/N_{dat} values, we refer to appendix C, where we also compare its performance to the DIS only fit and the EPPS16 nPDF set.

The inclusion of LHC data in the fit primarily affects the low x behaviour of the fits. For $x \lesssim 0.1$, the uncertainties are reduced quite dramatically and the nuclear shadowing effect at low x is now clearly visible for the valence and sea quarks. While the impact on central values is not as dramatic for $x \gtrsim 0.1$, there is a slight reduction in the uncertainties.

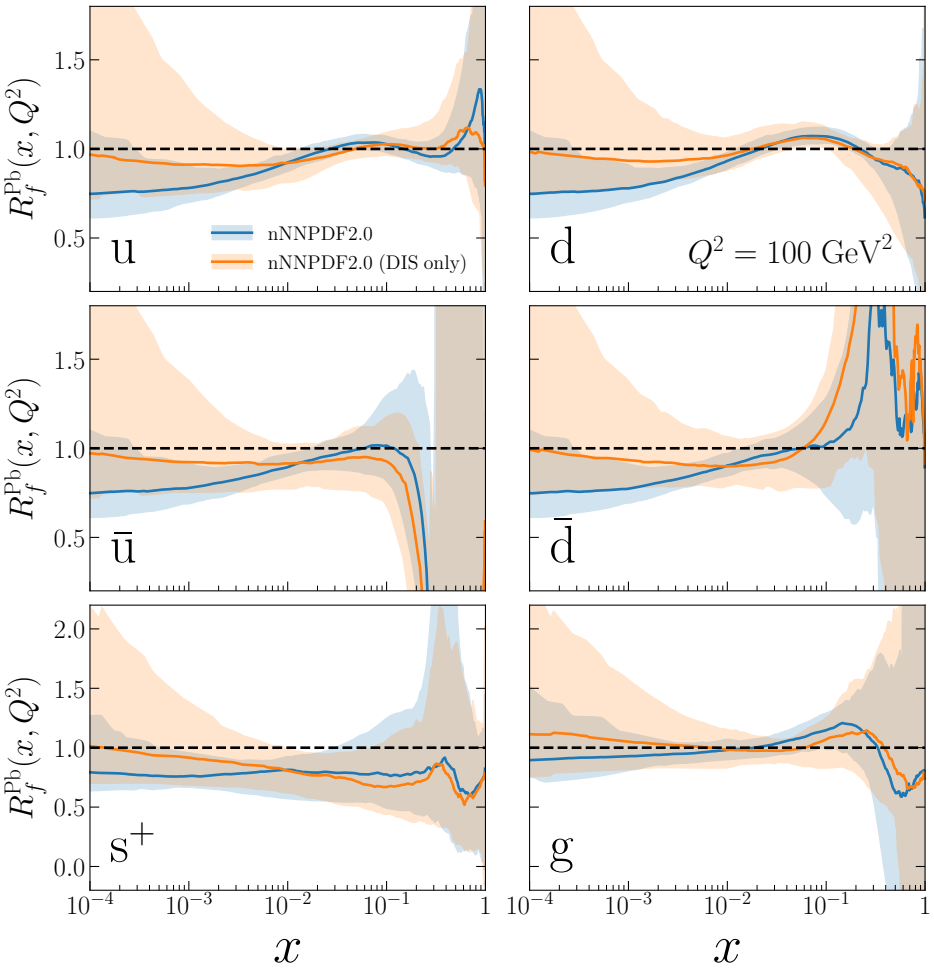

Figure 5.2: Nuclear modification factor for lead as determined by the DIS only fit and the global nNNPDF2.0 fit, normalised with respect to the free proton baseline. The shaded bands correspond to the 90% confidence levels. Note that the global fit exhibits a pronounced shadowing effect and decreased uncertainties. Figure adapted from (52).

In figure 5.3, we show the Pb ($A = 208$) nPDFs, at $Q^2 = 100\,\text{GeV}^2$, based on a 1000 replica fit. We show the valence u and d quarks, the $\bar{u}$, s and c sea quarks, and the gluon. Again, the shaded areas correspond to the 90% confidence band. Note the clear separation of the various flavours. These nPDFs were determined by applying the DGLAP evolution equations to the nPDFs fitted at energy $Q_0^2 = 1\,\text{GeV}^2$.

Now, we can use the global fit and examine how it performs on the CMS Z data. In the top panel of figure 5.4, we show the calculated cross-sections as compared to the experimental data. We show both the free proton ($A = 1$) fit and the lead ($A = 208$) fit. The middle panel shows the ratio of the data to the $A = 208$ fit and the lower panel shows the nuclear modification factor $R_A = f^{(N/A)}/f^{(N)}$. The data/theory ratio is close to one over the whole rapidity spectrum, indicating the nNNPDF2.0 fit yields an accurate prediction for this data, which is further validated by its $\chi^2/N_{dat} = 0.521$. The fit also shows a clear nuclear modification of up to $\sim 10\%$ in this rapidity range.

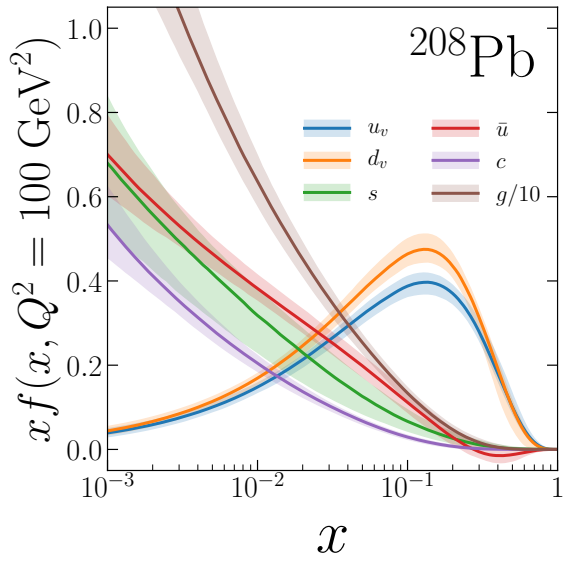

Figure 5.3: Nuclear PDFs for lead ($A = 208$) as determined by the nNNPDF2.0 global fit, evolved to $Q^2 = 100\,\mathrm{GeV}^2$ with the DGLAP equations. The shaded areas correspond to the 90% confidence bands. Note that the nNNPDF2.0 fit displays a clear quark flavour separation. Figure adapted from (52).

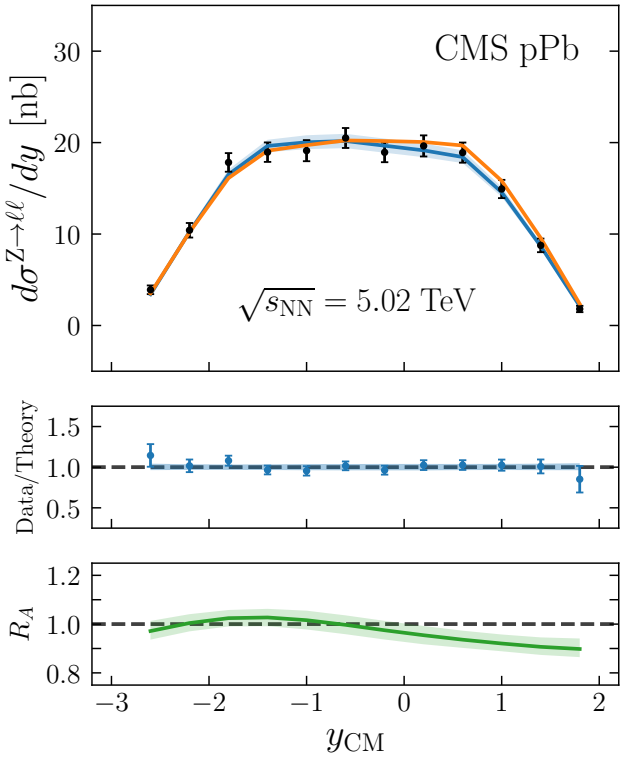

Figure 5.4: Differential cross-section for pPb collisions at $\sqrt{s_{NN}} = 5.12$ TeV. The orange line corresponds to the cross-section calculated with the nNNPDF2.0 proton ($A = 1$) PDF and the blue line corresponds to the lead ($A = 208$) PDF. The shaded band corresponds to the 90% confidence level. The middle and lower panel corresponds to the data/theory ratio for the $A = 208$ fit and the nuclear modification factor, respectively. Figure adapted from (52).

5.2 ATLAS photon production

In this experiment, inclusive, isolated, prompt photon production was studied for pPb collisions in the ATLAS detector (112). The experiment was performed at a centre-of-mass energy per nucleon of $\sqrt{s_{NN}} = 8.16$ TeV and an integrated luminosity of $\mathcal{L} = 165\,\text{nb}^{-1}$. The centre-of-mass pseudorapidity range is divided into three regions: $(-2.83, -2.02), (-1.84, 0.91)$, and $(1.09, 1.90)$ and the detected photons must have transverse energy $E_T^\gamma > 20$ GeV. In order for a photon to be identified as originating from an inclusive, prompt production process, it must fulfil the isolation requirements of $E_{iso}^\gamma < 4.8$ GeV $+4.2 \times 10^{-3} E_T^\gamma$ within a cone of $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} = 0.4$ around the photon. See also appendix B for the full settings used for the `APPLgrid` generation.

For constructing the theoretical predictions for the ATLAS data, we have used a patched version of the `MCFM v6.8` software with the same settings as reference (116). In this patch, the calculation of the experimental isolation conditions has been altered so that we do not need to calculate the fragmentation component. (52)

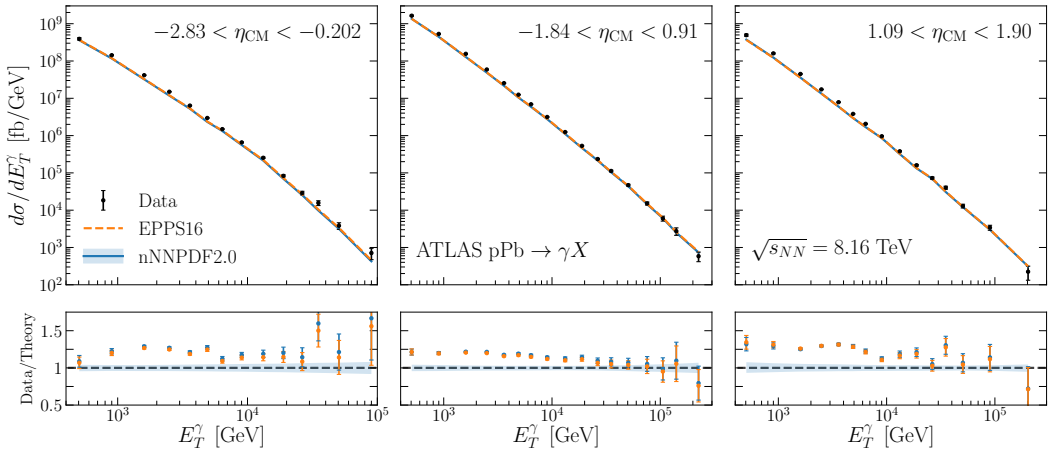

Figure 5.5: Cross-sections for the ATLAS photon production in the three rapidity bins, fitted with both the nNNPDF2.0 global fit and the EPPS16 NLO fit. The upper panel shows the absolute cross-sections and the lower panel shows the ratio of the data to the theory predictions. The two fits are in reasonable agreement with each other, but deviate significantly from the experimental data. Figure adapted from (52).

In figure 5.5, we show a fit of the global nNNPDF2.0 and EPPS16 (7) fits to the ATLAS photon data in the three rapidity bins. The upper panels show the fits to the absolute data and the lower panels show the ratio of the experimental data as normalised to the theoretical predictions. As can be seen in the ratio plots, although the nNNPDF2.0 and EPPS16 sets are in agreement with each other, they do not describe the data well. It should be noted that this same behaviour was (qualitatively) present in the original analysis done by the ATLAS collaboration.

Note that the calculation of the nNNPDF2.0 and EPPS16 sets are based on different Monte Carlo simulation algorithms. While nNNPDF2.0 uses `MCFM`,

EPPS16 is based on the JETPHOX (117) software. Despite the different software, there is a reasonable agreement between the two calculations. However, the theory calculations undershoot the data for nearly all datapoints. This disagreement is reflected in the χ^2 values, with the global nNNPDF2.0 fit having a $\chi^2/N_{dat} = 9.1$, 10.5, and 8.5 in the forward, central and backwards rapidity bins, respectively, with similar numbers for the EPPS16 fit. In appendix C, we have included the χ^2/N_{dat} values for EPPS16 on the data sets included in the nNNPDF2.0 global fit, where applicable, as a reference to their overall agreement across data sets.

In order to investigate the issue with the description of this data, it is instructive to include it in our nPDF fit. Thus, we must first implement the data in the buildmaster. The uncertainties for this experiment are presented individually, for each source of systematic uncertainty. From the reference it is not clear whether they are additive or multiplicative, so, as per the NNPDF policy, we treat all of them as multiplicative and correlated errors (apart from the statistical uncertainty). The purity and detector performance errors need to be symmetrised and the central value of the data is shifted accordingly.

Including the ATLAS photon data in the fit (unsurprisingly) yields better results ($\chi^2/N_{dat} = 6.1$, 7.5, and 5.7) than the global nNNPDF2.0 fit, but this is still far from satisfactory. The agreement between nNNPDF2.0 and EPPS16, and their mutual disagreement with this data, is extra puzzling because the NNPDF3.1 proton PDF is known to describe prompt photon production in the ATLAS detector well for pp collisions at both $\sqrt{s_{NN}} = 8$ TeV and $\sqrt{s_{NN}} = 13$ TeV (116). Until the origin of this data-theory discrepancy is fully understood, including ATLAS photon data in a global nPDF fit will be ineffective.

Chapter 6

Summary and outlook

The calculation of hadronic observables depends on non-perturbative objects called parton distribution functions, that govern the momentum of quarks and gluons within hadrons. In processes involving nuclei, the PDFs are further modified non-trivially, necessitating a separate determination of nuclear PDFs from experimental data.

In this thesis we have discussed the addition of CMS Z boson production data from pPb collisions into the NNPDF framework. This dataset was added to the nNNPDF2.0 nuclear PDF set, contributing to improved quark flavour separation over its predecessor nNNPDF1.0. The inclusion of this data and that of similar experiments also leads to a dramatic improvement of the nuclear modification displayed by the fit, as shown in figure 5.2. Overall, these results show the power of factorisation theorems to describe the nuclear modification of PDFs.

We have also presented a phenomenological exploration of the nNNPDF2.0 fit to prompt photon production data in pPb collisions in the ATLAS detector. Both the nNNPDF2.0 and EPPS16 nPDF sets do not describe this data well, see figure 5.5. Including the data in the training set does not improve the quality of the theory predictions for this data to a satisfactory level. This implies a further investigation of these processes is necessary, especially considering that the analysis of the ATLAS collaboration itself shows a similar, poor description and the fact that the same process in pp collisions is well described by free proton PDF sets.

The current nNNPDF2.0 PDF set displays relatively large uncertainties for the gluon. In order to remedy this, one could study prompt photon production data. However, we have seen that this might not actually improve the quality of the fit, if similar results are achieved as for the ATLAS photon data studied in this thesis. Alternatively, one could investigate the inclusion of pPb dijet production data in the next global fit. Dijet production in LHC run I pp collisions has been studied recently ([51]) at NNLO and it has been shown to constrain the gluon at large x. The corresponding pPb case has been shown to greatly affect the gluon nuclear modification ([118]) in an EPPS16-based analysis, indicating it would improve the nNNPDF2.0 fit as well.

Acknowledgements

Over the past year, I have had the opportunity to study a fascinating topic with a group of brilliant people. Here, I would like to express my utmost gratitude to those that helped and supported me during my time at Nikhef.

First and foremost, I would like to thank Dr. Juan Rojo for offering me this project and for welcoming me into his group. I also want to express my gratitude to Dr. Alexey Boyarsky for functioning as my LION supervisor.

In particular, I am immensely grateful to Rabah Khalek for assisting me in this project from start to finish, regardless of the amount of other, arguably more important, work he had to do. Without you, I would not have gotten nearly as far as I have.

I thank Jake Ethier and Emanuele Nocera for being ready to answer any questions I had on both physics and the NNPDF code.

I thank Ferran Faura Iglesias for the insightful discussions during our joint investigation of PDFs and Jaco ter Hoeve for the help in understanding QCD. I would also like to thank the Master's students in the Nikhef theory group for teaching me about their research during our plenary meetings.

Appendix A

Other free proton PDFs

MSTW08 PDF parametrisation

The DSSZ nuclear PDF set uses the MSTW08 ([79]) free proton PDF, which is parametrised at $Q_0^2 = 1\text{GeV}^2$ as:

$$xu_v = A_u x^{\eta_1}(1-x)^{\eta_2}(1+\varepsilon_u\sqrt{x}+\gamma_u x) \tag{A.1}$$
$$xd_v = A_d x^{\eta_3}(1-x)^{\eta_4}(1+\varepsilon_d\sqrt{x}+\gamma_d x) \tag{A.2}$$
$$xS = A_S x^{\delta_S}(1-x)^{\eta_S}(1+\varepsilon_S\sqrt{x}+\gamma_S x) \tag{A.3}$$
$$x\Delta = A_\Delta x^{\eta_\Delta}(1-x)^{\eta_S+2}(1+\gamma_\Delta x+\delta_\Delta x^2) \tag{A.4}$$
$$xg = A_g x^{\delta_g}(1-x)^{\eta_g}(1+\varepsilon_g\sqrt{x}+\gamma_g x) + A_{g'}x^{\delta_{g'}}(1-x)^{\eta_{g'}} \tag{A.5}$$
$$x(s+\bar{s}) = A_+ x^{\delta_S}(1-x)^{\eta_+}(1+\varepsilon_S\sqrt{x}+\gamma_S x) \tag{A.6}$$
$$x(s-\bar{s}) = A_- x^{\delta_-}(1-x)^{\eta_-}(1-x/x_0) \tag{A.7}$$

where $q_v = q - \bar{q}$, $\Delta = \bar{d} - \bar{u}$ and $S = 2(\bar{u}+\bar{d}) + s + \bar{s}$. Using the flavour and momentum sum rules, the values of A_g, A_u, A_d and x_0 can be expressed in terms of other parameters. In principle, there are then 30 free PDF parameters (including α_s), which is reduced to 28 due to strong (anti-)correlations between some of the parameters. When including the Hessian uncertainty calculation, this is extended to a total of 49 free parameters.

JR09 PDF parametrisation

The JR09 PDF ([80]) is used by the KA15 nPDF as the free proton prior at NNLO. They fit the $u_v, d_v, \Delta = \bar{d} - \bar{u}, \bar{d} + \bar{u}, s = \bar{s}$ and g PDFs at various values of Q_0, with their standard fit being at $Q_0^2 = 2\text{GeV}^2$. They parametrise the PDF interpolation functions with a simple polynomial in $\sqrt{x}$:

$$xf_i = N_i x^{a_i}(1-x)^{b_i}(1+A_i\sqrt{x}+B_i x) \tag{A.8}$$

Then, by setting $A_g = B_g = 0$, using $\bar{s} = s = (\bar{d}+\bar{u})/4$ and using the QCD flavour sum rule, this fit has a total of 21 free parameters.

CT14 PDF parametrisation

The CT14nlo ([78]) PDF set is used by EPPS16 as their proton baseline. They parametrise the $g, u, \bar{u}, d, \bar{d}, s$ PDFs with $s = \bar{s}$ at $Q_0 = 1.4$ GeV by using a fourth order polynomial in $y = \sqrt{x}$ as the interpolation function $\mathcal{I}$. Instead of fitting this function from the data, they instead transform it to a linear combination of Bernstein polynomials. For, e.g., u_v, this then becomes:

$$P_{u_v} = d_0 p_0(y) + d_1 p_1(y) + d_2 p_2(y) + d_3 p_3(y) + d_4 p_4(y) \tag{A.9}$$

where

$$p_0(y) = (1-y)^4 \tag{A.10}$$
$$p_1(y) = 4y(1-y)^3 \tag{A.11}$$
$$p_2(y) = 6y^2(1-y)^2 \tag{A.12}$$
$$p_3(y) = 4y^2(1-y) \tag{A.13}$$
$$p_4(y) = y^4 \tag{A.14}$$

The d_k parameters are then fitted from the data and the interpolation function can then be calculated by reverting P_{u_v} back to its simple polynomial shape:

$$P_{u_v} = c_0 + c_1 y + c_2 y^2 + c_3 y^3 + c_4 y^4 \tag{A.15}$$

In practice, not all d_k parameters are free parameters. The value for d_4 is set to 1 and supplanted with an overall constant factor, determined by the flavour sum rule $\int_0^1 dx u_v = 2$. Also, to suppress deviations from the high-x $(1-x)^{\beta_{u_v}}$ behaviour, $d_3 = 1 + \alpha_{u_v}/2$ is set. The effective exponents for d_v are also set to be equal to those of u_v. Ultimately, this leaves us with a total of 28 free parameters.

Appendix B

MCFM settings

CMS Z production

```
'6.8'[file version number]

[Flags to specify the mode in which MCFM is run]
-1 [nevtrequested]
.false. [creatent]
.false. [skipnt]
.false. [dswhisto]
.true.          [creategrid]
.false. [writetop]
.false. [writedat]
.false. [writegnu]
.false. [writeroot]
.false. [writepwg]

[General options to specify the process and execution]
31   [nproc]
'tota'   [part 'lord','real' or 'virt','tota']
'CMSpPbZ5TEV'['runstring']
5020d0 [sqrts in GeV]
+1 [ih1 =1 for proton and -1 for antiproton]
+1 [ih2 =1 for proton and -1 for antiproton]
125.09d0 [hmass]
91.1876d0 [scale:QCD scale choice]
91.1876d0 [facscale:QCD fac_scale choice]
'no'[dynamicscale]
.false. [zerowidth]
.false. [removebr]
10 [itmx1, number of iterations for pre-conditioning]
10000 [ncall1]
10 [itmx2, number of iterations for final run]
200000 [ncall2]
1089 [ij]
.false. [dryrun]
```

```
.true. [Qflag]
.true. [Gflag]

[Heavy quark masses]
173.1d0 [top mass]
4.18d0 [bottom mass]
1.28d0 [charm mass]

[Pdf selection]
'CT10.00'[pdlabel]
4 [NGROUP, see PDFLIB]
46 [NSET - see PDFLIB]
CT10nlo.LHgrid [LHAPDF group]
0            [LHAPDF set]

[Jet definition and event cuts]
60d0 [m34min]
120d0 [m34max]
0d0 [m56min]
14000d0 [m56max]
.true. [inclusive]
'ankt'[algorithm]
120d0 [ptjet_min]
0d0 [|etajet|_min]
3d0 [|etajet|_max]
0.3d0 [Rcut_jet]
.true. [makecuts]
20d0 [ptlepton_min]
-2.865d0,1.935d0 [|etalepton|_max]
0d0,0d0         [|etalepton|_veto]
0d0 [ptmin_missing]
20d0 [ptlepton(2nd+)_min]
-2.865d0,1.935d0 [|etalepton(2nd+)|_max]
0d0,0d0         [|etalepton(2nd+)|_veto]
0d0 [minimum (3,4) transverse mass]
0d0 [R(jet,lept)_min]
0d0 [R(lept,lept)_min]
0d0 [Delta_eta(jet,jet)_min]
.false. [jets_opphem]
0 [lepbtwnjets_scheme]
0d0             [ptmin_bjet]
99d0 [etamax_bjet]

[Settings for photon processes]
.false.         [fragmentation included]
'GdRG__LO'      [fragmentation set]
80d0        [fragmentation scale]
20d0 [ptmin_photon]
```

```
2.5d0 [etamax_photon]
20d0 [ptmin_photon(2nd)]
20d0 [ptmin_photon(3rd)]
0d0           [R(photon,lept)_min]
0.4d0           [R(photon,photon)_min]
0.4d0           [R(photon,jet)_min]
0.4d0           [cone size for isolation]
0.5d0 [epsilon_h, energy fraction for isolation]

[Anomalous couplings of the W and Z]
0.0d0 [Delta_g1(Z)]
0.0d0 [Delta_K(Z)]
0.0d0 [Delta_K(gamma)]
0.0d0 [Lambda(Z)]
0.0d0 [Lambda(gamma)]
0.0d0 [h1(Z)]
0.0d0 [h1(gamma)]
0.0d0 [h2(Z)]
0.0d0 [h2(gamma)]
0.0d0 [h3(Z)]
0.0d0 [h3(gamma)]
0.0d0 [h4(Z)]
0.0d0 [h4(gamma)]
2.0d0 [Form-factor scale, in TeV]

[Anomalous width of the Higgs]
1d0             [Gamma_H/Gamma_H(SM)]

[How to resume/save a run]
.false. [readin]
.false. [writeout]
''[ingridfile]
''[outgridfile]

[Technical parameters that should not normally be changed]
.false. [debug]
.true. [verbose]
.false. [new_pspace]
.false. [virtonly]
.false. [realonly]
.true. [spira]
.false. [noglue]
.false. [ggonly]
.false. [gqonly]
.false. [omitgg]
.false. [vanillafiles]
1 [nmin]
2 [nmax]
```

```
.true. [clustering]
.false. [realwt]
0 [colourchoice]
1d-2 [rtsmin]
1d-4 [cutoff]
0.1d0 [aii]
0.1d0 [aif]
0.1d0 [afi]
1d0 [aff]
1d0             [bfi]
1d0             [bff]
```

ATLAS prompt photon production

```
'6.8'[file version number]

[Flags to specify the mode in which MCFM is run]
-1 [nevtrequested]
.false. [creatent]
.false. [skipnt]
.false. [dswhisto]
.true.          [creategrid]
.false. [writetop]
.false. [writedat]
.false. [writegnu]
.false. [writeroot]
.false. [writepwg]

[General options to specify the process and execution]
280   [nproc]
'tota'   [part 'lord','real' or 'virt','tota']
'nATLAS_pPb_PHT_8TEV'['runstring']
8160d0 [sqrts in GeV]
+1 [ih1 =1 for proton and -1 for antiproton]
+1 [ih2 =1 for proton and -1 for antiproton]
125.1d0 [hmass]
1d0 [scale:QCD scale choice]
1d0 [facscale:QCD fac_scale choice]
'pt(photon)'[dynamicscale]
.false. [zerowidth]
.false. [removebr]
10 [itmx1, number of iterations for pre-conditioning]
10000 [ncall1]
15 [itmx2, number of iterations for final run]
200000 [ncall2]
1089 [ij]
.false. [dryrun]
.true. [Qflag]
.true. [Gflag]

[Heavy quark masses]
172.9d0 [top mass]
4.18d0 [bottom mass]
1.27d0 [charm mass]

[Pdf selection]
'mstw8nl'[pdlabel]
4 [NGROUP, see PDFLIB]
46 [NSET - see PDFLIB]
NNPDF31_nlo_as_0118.LHgrid [LHAPDF group]
```

```
0            [LHAPDF set]

[Jet definition and event cuts]
0d0 [m34min]
14000d0 [m34max]
0d0 [m56min]
14000d0 [m56max]
.true. [inclusive]
'ankt'[algorithm]
0d0 [ptjet_min]
0d0 [|etajet|_min]
99d0 [|etajet|_max]
0.2d0 [Rcut_jet]
.false. [makecuts]
0d0 [ptlepton_min]
99d0 [|etalepton|_max]
0d0,0d0         [|etalepton|_veto]
0d0 [ptmin_missing]
0d0 [ptlepton(2nd+)_min]
99d0 [|etalepton(2nd+)|_max]
0d0,0d0         [|etalepton(2nd+)|_veto]
0d0 [minimum (3,4) transverse mass]
0d0 [R(jet,lept)_min]
0d0 [R(lept,lept)_min]
0d0 [Delta_eta(jet,jet)_min]
.false. [jets_opphem]
0 [lepbtwnjets_scheme]
0d0             [ptmin_bjet]
99d0 [etamax_bjet]

[Settings for photon processes]
.false.         [fragmentation included]
'GdRG__LO'      [fragmentation set]
80d0        [fragmentation scale]
20d0 [ptmin_photon]
2.83d0 [etamax_photon]
0d0 [ptmin_photon(2nd)]
0d0 [ptmin_photon(3rd)]
0d0           [R(photon,lept)_min]
0d0           [R(photon,photon)_min]
0d0           [R(photon,jet)_min]
0.4d0           [cone size for isolation]
0.1d0 [epsilon_h, energy fraction for isolation]

[Anomalous couplings of the W and Z]
0.0d0 [Delta_g1(Z)]
0.0d0 [Delta_K(Z)]
0.0d0 [Delta_K(gamma)]
```

```
0.0d0 [Lambda(Z)]
0.0d0 [Lambda(gamma)]
0.0d0 [h1(Z)]
0.0d0 [h1(gamma)]
0.0d0 [h2(Z)]
0.0d0 [h2(gamma)]
0.0d0 [h3(Z)]
0.0d0 [h3(gamma)]
0.0d0 [h4(Z)]
0.0d0 [h4(gamma)]
2.0d0 [Form-factor scale, in TeV]

[Anomalous width of the Higgs]
1d0             [Gamma_H/Gamma_H(SM)]

[How to resume/save a run]
.false. [readin]
.false. [writeout]
''[ingridfile]
''[outgridfile]

[Technical parameters that should not normally be changed]
.false. [debug]
.true. [verbose]
.true. [new_pspace]
.false. [virtonly]
.false. [realonly]
.true. [spira]
.false. [noglue]
.false. [ggonly]
.false. [gqonly]
.false. [omitgg]
.false. [vanillafiles]
1 [nmin]
2 [nmax]
.true. [clustering]
.false. [realwt]
0 [colourchoice]
1d-2 [rtsmin]
1d-4 [cutoff]
0.1d0 [aii]
0.1d0 [aif]
0.1d0 [afi]
1d0 [aff]
1d0             [bfi]
1d0             [bff]
```

Appendix C

χ^2 tables

		nNNPDF2.0 (DIS)	nNNPDF2.0	EPPS16nlo
Dataset	$N_{\rm dat}$	$\chi^2/N_{\rm dat}$	$\chi^2/N_{\rm dat}$	$\chi^2/N_{\rm dat}$
NMC (He/D)	13	1.11	1.129	0.829
SLAC (He/D)	3	0.623	0.638	0.152
NMC (Li/D)	12	1.083	1.166	0.74
SLAC (Be/D)	3	1.579	1.719	0.098
EMC (C/D)	12	1.292	1.321	1.174
FNAL (C/D)	3	0.932	0.838	0.985
NMC (C/D)	26	2.002	2.171	0.872
SLAC (C/D)	2	0.286	0.251	1.075
BCDMS (N/D)	9	2.439	2.635	n/a
SLAC (Al/D)	3	0.606	0.864	0.326
EMC (Ca/D)	3	1.72	1.722	1.82
FNAL (Ca/D)	3	1.253	1.194	1.354
NMC (Ca/D)	12	1.503	1.747	1.772
SLAC (Ca/D)	2	0.82	0.771	1.642
BCDMS (Fe/D)	16	2.244	2.743	0.765
EMC (Fe/D)	58	0.827	0.875	0.445
SLAC (Fe/D)	8	2.171	2.455	1.06
EMC (Cu/D)	27	0.523	0.572	0.714
SLAC (Ag/D)	2	0.667	0.691	1.595
EMC (Sn/D)	8	2.197	2.248	2.265
FNAL (Xe/D)	4	0.414	0.384	n/a
SLAC (Au/D)	3	1.216	1.353	1.916
FNAL (Pb/D)	3	2.243	2.168	2.044
NMC (Be/C)	14	0.268	0.269	0.27
NMC (C/Li)	9	1.063	1.117	0.9
NMC (Al/C)	14	0.345	0.354	0.396
NMC (Ca/C)	23	0.468	0.44	0.564
NMC (Fe/C)	14	0.663	0.667	0.751
NMC (Sn/C)	119	0.607	0.638	0.626
NMC (Ca/Li)	9	0.259	0.276	0.15

Table C.1: The values of the χ^2 per data point for the DIS neutral current structure function datasets included in nNNPDF2.0. We compare the χ^2/N_{dat} of the nNNPDF2.0 DIS-only fit with those obtained by the global nNNPDF2.0 fit and EPPS16. Table adapted from (52)

		nNNPDF2.0 (DIS)	nNNPDF2.0	EPPS16nlo
Dataset	$N_{\rm dat}$	$\chi^2/N_{\rm dat}$	$\chi^2/N_{\rm dat}$	$\chi^2/N_{\rm dat}$
NuTeV ($\bar{\nu}$Fe)	37	0.946	1.094	*0.639*
NuTeV (νFe)	39	0.287	0.264	*0.381*
CHORUS ($\bar{\nu}$Pb)	423	0.938	0.97	1.107
CHORUS (νPb)	423	1.007	1.015	1.024
ATLAS$^{\rm 5TEV}$ Z	14	*1.469*	1.134	1.12
CMS$^{\rm 5TeV}$ W$^-$	10	*1.688*	1.078	0.857
CMS$^{\rm 8TeV}$ W$^-$	24	*1.453*	0.72	*0.825*
CMS$^{\rm 5TeV}$ W$^+$	10	*2.32*	1.125	1.211
CMS$^{\rm 8TeV}$ W$^+$	24	*3.622*	0.772	*0.951*
CMS$^{\rm 5TeV}$ Z	12	*0.58*	0.52	0.639
Total	**1467**	**1.013**	**0.976**	**0.896**

Table C.2: Same as Table C.1 now for the datasets newly included in nNNPDF2.0: charged current DIS structure functions and gauge boson production at the LHC. We also provide the values of $\chi^2/N_{\rm dat}$ for the total dataset. Values in italics indicate predictions for datasets not included in the corresponding fit. Table adapted from (52)

References

[1] J. Rojo, *The Partonic Content of Nucleons and Nuclei*, 1910.03408.

[2] J. Gao, L. Harland-Lang and J. Rojo, *The Structure of the Proton in the LHC Precision Era*, *Phys. Rept.* **742** (2018) 1 [1709.04922].

[3] J. J. Ethier and E. R. Nocera, *Parton Distributions in Nucleons and Nuclei*, *Ann. Rev. Nucl. Part. Sci.* (2020) 1 [2001.07722].

[4] D. de Florian, R. Sassot, P. Zurita and M. Stratmann, *Global Analysis of Nuclear Parton Distributions*, *Phys. Rev. D* **85** (2012) 074028 [1112.6324].

[5] H. Khanpour and S. Atashbar Tehrani, *Global Analysis of Nuclear Parton Distribution Functions and Their Uncertainties at Next-to-Next-to-Leading Order*, *Phys. Rev. D* **93** (2016) 014026 [1601.00939].

[6] K. Kovarik et al., *nCTEQ15 - Global analysis of nuclear parton distributions with uncertainties in the CTEQ framework*, *Phys. Rev. D* **93** (2016) 085037 [1509.00792].

[7] K. J. Eskola, P. Paakkinen, H. Paukkunen and C. A. Salgado, *EPPS16: Nuclear parton distributions with LHC data*, *Eur. Phys. J. C* **77** (2017) 163 [1612.05741].

[8] M. Walt, I. Helenius and W. Vogelsang, *A QCD analysis for nuclear PDFs at NNLO*, *PoS* **DIS2019** (2019) 039 [1908.04983].

[9] M. Walt, I. Helenius and W. Vogelsang, *Open-source QCD analysis of nuclear parton distribution functions at NLO and NNLO*, *Phys. Rev. D* **100** (2019) 096015 [1908.03355].

[10] J. Rojo, *The neural network approach to parton distribution functions*, Ph.D. thesis, 2006. hep-ph/0607122.

[11] S. Forte, L. Garrido, J. I. Latorre and A. Piccione, *Neural network parametrization of deep inelastic structure functions*, *JHEP* **05** (2002) 062 [hep-ph/0204232].

[12] S. Forte, J. I. Latorre, L. Magnea and A. Piccione, *Determination of $\alpha(s)$ from scaling violations of truncated moments of structure functions*, *Nucl. Phys. B* **643** (2002) 477 [hep-ph/0205286].

[13] J. Rojo and J. I. Latorre, *Neural network parametrization of spectral functions from hadronic tau decays and determination of QCD vacuum condensates*, *JHEP* **01** (2004) 055 [hep-ph/0401047].

[14] NNPDF collaboration, *Unbiased determination of the proton structure function F_2^p with faithful uncertainty estimation*, *JHEP* **03** (2005) 080 [hep-ph/0501067].

[15] NNPDF collaboration, *Neural network determination of parton distributions: The Nonsinglet case*, *JHEP* **03** (2007) 039 [hep-ph/0701127].

[16] NNPDF collaboration, *A Determination of parton distributions with faithful uncertainty estimation*, *Nucl. Phys. B* **809** (2009) 1 [0808.1231].

[17] L. Del Debbio, A. Guffanti and A. Piccione, *The Bjorken sum rule with Monte Carlo and Neural Network techniques*, *JHEP* **11** (2009) 060 [0907.2506].

[18] NNPDF collaboration, *Precision determination of electroweak parameters and the strange content of the proton from neutrino deep-inelastic scattering*, *Nucl. Phys. B* **823** (2009) 195 [0906.1958].

[19] NNPDF collaboration, *Fitting Parton Distribution Data with Multiplicative Normalization Uncertainties*, *JHEP* **05** (2010) 075 [0912.2276].

[20] R. D. Ball, L. Del Debbio, S. Forte, A. Guffanti, J. I. Latorre, J. Rojo et al., *A first unbiased global NLO determination of parton distributions and their uncertainties*, *Nucl. Phys. B* **838** (2010) 136 [1002.4407].

[21] NNPDF collaboration, *Reweighting NNPDFs: the W lepton asymmetry*, *Nucl. Phys. B* **849** (2011) 112 [1012.0836].

[22] R. D. Ball, V. Bertone, F. Cerutti, L. Del Debbio, S. Forte, A. Guffanti et al., *Impact of Heavy Quark Masses on Parton Distributions and LHC Phenomenology*, *Nucl. Phys. B* **849** (2011) 296 [1101.1300].

[23] NNPDF collaboration, *On the Impact of NMC Data on NLO and NNLO Parton Distributions and Higgs Production at the Tevatron and the LHC*, *Phys. Lett. B* **704** (2011) 36 [1102.3182].

[24] S. Lionetti, R. D. Ball, V. Bertone, F. Cerutti, L. Del Debbio, S. Forte et al., *Precision determination of α_s using an unbiased global NLO parton set*, *Phys. Lett. B* **701** (2011) 346 [1103.2369].

[25] NNPDF collaboration, *Unbiased global determination of parton distributions and their uncertainties at NNLO and at LO*, *Nucl. Phys. B* **855** (2012) 153 [1107.2652].

[26] R. D. Ball, V. Bertone, F. Cerutti, L. Del Debbio, S. Forte, A. Guffanti et al., *Reweighting and Unweighting of Parton Distributions and the LHC W lepton asymmetry data*, 2012. 10.1016/j.nuclphysb.2011.10.018.

[27] R. D. Ball, V. Bertone, L. Del Debbio, S. Forte, A. Guffanti, J. I. Latorre et al., *Precision NNLO determination of alpha_s(M_Z) using an unbiased global parton set*, *Phys. Lett. B* **707** (2012) 66 [1110.2483].

[28] R. D. Ball et al., *Parton distributions with LHC data*, *Nucl. Phys. B* **867** (2013) 244 [1207.1303].

[29] NNPDF collaboration, R. D. Ball, V. Bertone, L. Del Debbio, S. Forte, A. Guffanti, J. Rojo et al., *Theoretical issues in PDF determination and associated uncertainties*, 2013. 10.1016/j.physletb.2013.05.019.

[30] NNPDF collaboration, R. D. Ball, S. Forte, A. Guffanti, E. R. Nocera, G. Ridolfi and J. Rojo, *Unbiased determination of polarized parton distributions and their uncertainties*, 2013. 10.1016/j.nuclphysb.2013.05.007.

[31] NNPDF collaboration, R. D. Ball, V. Bertone, S. Carrazza, L. Del Debbio, S. Forte, A. Guffanti et al., *Parton distributions with QED corrections*, 2013. 10.1016/j.nuclphysb.2013.10.010.

[32] NNPDF collaboration, *Polarized Parton Distributions at an Electron-Ion Collider*, *Phys. Lett. B* **728** (2014) 524 [1310.0461].

[33] NNPDF collaboration, *A first unbiased global determination of polarized PDFs and their uncertainties*, *Nucl. Phys. B* **887** (2014) 276 [1406.5539].

[34] NNPDF collaboration, *Parton distributions for the LHC Run II*, *JHEP* **04** (2015) 040 [1410.8849].

[35] M. Bonvini, S. Marzani, J. Rojo, L. Rottoli, M. Ubiali, R. D. Ball et al., *Parton distributions with threshold resummation*, *JHEP* **09** (2015) 191 [1507.01006].

[36] NNPDF collaboration, R. D. Ball, V. Bertone, M. Bonvini, S. Carrazza, S. Forte, A. Guffanti et al., *A Determination of the Charm Content of the Proton*, 2016. 10.1140/epjc/s10052-016-4469-y.

[37] NNPDF collaboration, R. D. Ball et al., *Parton distributions from high-precision collider data*, 2017. 10.1140/epjc/s10052-017-5199-5.

[38] NNPDF collaboration, *A determination of the fragmentation functions of pions, kaons, and protons with faithful uncertainties*, *Eur. Phys. J. C* **77** (2017) 516 [1706.07049].

[39] R. D. Ball, V. Bertone, M. Bonvini, S. Marzani, J. Rojo and L. Rottoli, *Parton distributions with small-x resummation: evidence for BFKL dynamics in HERA data*, *Eur. Phys. J. C* **78** (2018) 321 [1710.05935].

[40] NNPDF collaboration, *Illuminating the photon content of the proton within a global PDF analysis*, *SciPost Phys.* **5** (2018) 008 [1712.07053].

[41] NNPDF collaboration, *Precision determination of the strong coupling constant within a global PDF analysis*, *Eur. Phys. J. C* **78** (2018) 408 [1802.03398].

[42] *Les Houches 2017: Physics at TeV Colliders Standard Model Working Group Report*, 3, 2018.

[43] NNPDF collaboration, *Charged hadron fragmentation functions from collider data*, *Eur. Phys. J. C* **78** (2018) 651 [1807.03310].

[44] NNPDF collaboration, *Nuclear Uncertainties in the Determination of Proton PDFs*, *Eur. Phys. J. C* **79** (2019) 282 [1812.09074].

[45] NNPDF collaboration, *Nuclear parton distributions from lepton-nucleus scattering and the impact of an electron-ion collider*, *Eur. Phys. J. C* **79** (2019) 471 [1904.00018].

[46] NNPDF collaboration, *A first determination of parton distributions with theoretical uncertainties*, *Eur. Phys. J.* **C** (2019) 79:838 [1905.04311].

[47] NNPDF collaboration, *Parton Distributions with Theory Uncertainties: General Formalism and First Phenomenological Studies*, *Eur. Phys. J. C* **79** (2019) 931 [1906.10698].

[48] E. R. Nocera, M. Ubiali and C. Voisey, *Single Top Production in PDF fits*, *JHEP* **05** (2020) 067 [1912.09543].

[49] S. Forte and Z. Kassabov, *Why α_s cannot be determined from hadronic processes without simultaneously determining the parton distributions*, *Eur. Phys. J. C* **80** (2020) 182 [2001.04986].

[50] S. Amoroso et al., *Les Houches 2019: Physics at TeV Colliders: Standard Model Working Group Report*, in *11th Les Houches Workshop on Physics at TeV Colliders: PhysTeV Les Houches*, 3, 2020, 2003.01700.

[51] R. Abdul Khalek et al., *Phenomenology of NNLO jet production at the LHC and its impact on parton distributions*, 2005.11327.

[52] R. Abdul Khalek, J. J. Ethier, J. Rojo and G. van Weelden, *nNNPDF2.0: Quark Flavor Separation in Nuclei from LHC Data*, 2006.14629.

[53] M. Gell-Mann, *A Schematic Model of Baryons and Mesons*, *Phys. Lett.* **8** (1964) 214.

[54] G. Zweig, *An SU(3) model for strong interaction symmetry and its breaking. Version 2*, .

[55] M. E. Peskin and D. V. Schroeder, *An Introduction to quantum field theory*. Addison-Wesley, Reading, USA, 1995.

[56] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, *Prog. Theor. Phys.* **49** (1973) 652.

[57] D. J. Gross and F. Wilczek, *Ultraviolet Behavior of Nonabelian Gauge Theories*, *Phys. Rev. Lett.* **30** (1973) 1343.

[58] H. Politzer, *Reliable Perturbative Results for Strong Interactions?*, *Phys. Rev. Lett.* **30** (1973) 1346.

[59] R. Ellis, W. Stirling and B. Webber, *QCD and collider physics*, vol. 8. Cambridge University Press, 2, 2011.

[60] S. Drell and T.-M. Yan, *Massive Lepton Pair Production in Hadron-Hadron Collisions at High-Energies*, *Phys. Rev. Lett.* **25** (1970) 316.

[61] J. C. Collins, D. E. Soper and G. F. Sterman, *Factorization of Hard Processes in QCD*, vol. 5, pp. 1–91, (1989), hep-ph/0409313, DOI.

[62] J. M. Campbell, J. Huston and W. Stirling, *Hard Interactions of Quarks and Gluons: A Primer for LHC Physics*, *Rept. Prog. Phys.* **70** (2007) 89 [hep-ph/0611148].

[63] Y. L. Dokshitzer, *Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics.*, *Sov. Phys. JETP* **46** (1977) 641.

[64] V. Gribov and L. Lipatov, *Deep inelastic e p scattering in perturbation theory*, *Sov. J. Nucl. Phys.* **15** (1972) 438.

[65] G. Altarelli and G. Parisi, *Asymptotic Freedom in Parton Language*, *Nucl. Phys. B* **126** (1977) 298.

[66] M. Botje, *QCDNUM: Fast QCD Evolution and Convolution*, *Comput. Phys. Commun.* **182** (2011) 490 [1005.1481].

[67] S. Malace, D. Gaskell, D. W. Higinbotham and I. Cloet, *The Challenge of the EMC Effect: existing data and future directions*, *Int. J. Mod. Phys. E* **23** (2014) 1430013 [1405.1270].

[68] European Muon collaboration, *The ratio of the nucleon structure functions F_2^N for iron and deuterium*, *Phys. Lett. B* **123** (1983) 275.

[69] N. Armesto, *Nuclear shadowing*, *J. Phys. G* **32** (2006) R367 [hep-ph/0604108].

[70] K. Eskola, V. Kolhinen and P. Ruuskanen, *Scale evolution of nuclear parton distributions*, *Nucl. Phys. B* **535** (1998) 351 [hep-ph/9802350].

[71] M. Hirai, S. Kumano and M. Miyama, *Determination of nuclear parton distributions*, *Phys. Rev. D* **64** (2001) 034003 [hep-ph/0103208].

[72] D. de Florian and R. Sassot, *Nuclear parton distributions at next-to-leading order*, *Phys. Rev. D* **69** (2004) 074028 [hep-ph/0311227].

[73] P. Paakkinen, *Nuclear parton distribution functions*, *Frascati Phys. Ser.* (2017) 33 [1802.05927].

[74] A. Kusina, F. Lyonnet, D. Clark, E. Godat, T. Jezo, K. Kovarik et al., *Vector boson production in pPb and PbPb collisions at the LHC and its impact on nCTEQ15 PDFs*, *Eur. Phys. J. C* **77** (2017) 488 [1610.02925].

[75] K. J. Eskola, P. Paakkinen and H. Paukkunen, *Non-quadratic improved Hessian PDF reweighting and application to CMS dijet measurements at 5.02 TeV*, *Eur. Phys. J. C* **79** (2019) 511 [1903.09832].

[76] K. J. Eskola, I. Helenius, P. Paakkinen and H. Paukkunen, *A QCD analysis of LHCb D-meson data in p+Pb collisions*, *JHEP* **05** (2020) 037 [1906.02512].

[77] L. Van Hove, *THEORETICAL PREDICTION OF A NEW STATE OF MATTER, THE 'QUARK - GLUON PLASMA' (ALSO CALLED 'QUARK MATTER')*, in *17th International Symposium on Multiparticle Dynamics*, pp. 801–818, 1986.

[78] S. Dulat, T.-J. Hou, J. Gao, M. Guzzi, J. Huston, P. Nadolsky et al., *New parton distribution functions from a global analysis of quantum chromodynamics*, *Phys. Rev. D* **93** (2016) 033006 [1506.07443].

[79] A. Martin, W. Stirling, R. Thorne and G. Watt, *Parton distributions for the LHC*, *Eur. Phys. J. C* **63** (2009) 189 [0901.0002].

[80] P. Jimenez-Delgado and E. Reya, *Dynamical NNLO parton distributions*, *Phys. Rev. D* **79** (2009) 074023 [0810.4274].

[81] T. Regge, *Introduction to complex orbital momenta*, *Nuovo Cim.* **14** (1959) 951.

[82] S. J. Brodsky and G. R. Farrar, *Scaling Laws at Large Transverse Momentum*, *Phys. Rev. Lett.* **31** (1973) 1153.

[83] R. D. Ball, E. R. Nocera and J. Rojo, *The asymptotic behaviour of parton distributions at small and large x*, *Eur. Phys. J. C* **76** (2016) 383 [1604.00024].

[84] J. Pumplin, D. Stump, R. Brock, D. Casey, J. Huston, J. Kalk et al., *Uncertainties of predictions from parton distribution functions. 2. The Hessian method*, *Phys. Rev. D* **65** (2001) 014013 [hep-ph/0101032].

[85] A. Buckley, J. Ferrando, S. Lloyd, K. Nordström, B. Page, M. Rüfenacht et al., *LHAPDF6: parton density access in the LHC precision era*, *Eur. Phys. J. C* **75** (2015) 132 [1412.7420].

[86] M. A. Shifman, A. Vainshtein and V. I. Zakharov, *QCD and Resonance Physics. Theoretical Foundations*, *Nucl. Phys. B* **147** (1979) 385.

[87] S. J. Brodsky, I. Schmidt and S. Liuti, *Is the Momentum Sum Rule Valid for Nuclear Structure Functions ?*, 8, 2019.

[88] W. S. McCulloch and W. Pitts, *A logical calculus of the ideas immanent in nervous activity*, *Bulletin of Mathematical Biology* **5** (1990) 115 .

[89] M. Minsky, *Perceptrons : an introduction to computational geometry*. MIT Press, Cambridge, MA [etc.], 1969.

[90] S. Carrazza, *Machine learning challenges in theoretical HEP*, *J. Phys. Conf. Ser.* **1085** (2018) 022003 [1711.10840].

[91] P. Mehta, M. Bukov, C.-H. Wang, A. G. Day, C. Richardson, C. K. Fisher et al., *A high-bias, low-variance introduction to machine learning for physicists*, *Physics Reports* **810** (2019) 1–124 [1803.08823].

[92] I. Goodfellow, Y. Bengio and A. Courville, *Deep Learning*. MIT Press, 2016.

[93] Y. Lecun, L. Bottou, Y. Bengio and P. Haffner, *Gradient-based learning applied to document recognition*, *Proceedings of the IEEE* **86** (1998) 2278.

[94] Q. V. Le, M. Ranzato, R. Monga, M. Devin, K. Chen, G. S. Corrado et al., *Building high-level features using large scale unsupervised learning*, 2011.

[95] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. Van Den Driessche et al., *Mastering the game of go with deep neural networks and tree search*, *nature* **529** (2016) 484.

[96] "Building a simple neural network with keras and tensorflow." https://whyaxis.me/2017/09/14/building-a-simple-neural-network-with-keras-and-tensorflow/.

[97] "Quora.com." https://www.quora.com/Which-signals-do-indicate-that-the-convolutional-neural-network-is-overfitted.

[98] B. Polyak, *Some methods of speeding up the convergence of iteration methods*, *USSR Computational Mathematics and Mathematical Physics* **4** (1964) 1 .

[99] Y. Nesterov, *A method for unconstrained convex minimization problem with the rate of convergence* $o(1/k^2)$, *Doklady AN USSR* **269** (1983) 543.

[100] S. Ruder, *An overview of gradient descent optimization algorithms*, 2016.

[101] J. Duchi, E. Hazan and Y. Singer, *Adaptive subgradient methods for online learning and stochastic optimization*, *Journal of Machine Learning Research* **12** (2011) 2121.

[102] G. Hinton, *Neural networks for machine learning*, 2012.

[103] X. Glorot and Y. Bengio, *Understanding the difficulty of training deep feedforward neural networks*, in *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*, Y. W. Teh and M. Titterington, eds., vol. 9 of *Proceedings of Machine Learning Research*, (Chia Laguna Resort, Sardinia, Italy), pp. 249–256, PMLR, 13–15 May, 2010, http://proceedings.mlr.press/v9/glorot10a.html.

[104] J. M. Campbell and R. Ellis, *An Update on vector boson pair production at hadron colliders*, *Phys. Rev. D* **60** (1999) 113006 [hep-ph/9905386].

[105] J. M. Campbell, R. Ellis and C. Williams, *Vector boson pair production at the LHC*, *JHEP* **07** (2011) 018 [1105.0020].

[106] J. M. Campbell, R. K. Ellis and W. T. Giele, *A Multi-Threaded Version of MCFM*, *Eur. Phys. J. C* **75** (2015) 246 [1503.06182].

[107] J. M. Campbell, R. K. Ellis and C. Williams, *MCFM v6.8 A Monte Carlo for FeMtobarn processes at Hadron Colliders Users Guide*, tech. rep., Fermilab, 2014.

[108] T. Carli, D. Clements, A. Cooper-Sarkar, C. Gwenlan, G. P. Salam, F. Siegert et al., *A posteriori inclusion of parton density functions in NLO QCD final-state calculations at hadron colliders: The APPLGRID Project*, *Eur. Phys. J. C* **66** (2010) 503 [0911.2985].

[109] E. Maguire, L. Heinrich and G. Watt, *HEPData: a repository for high energy physics data*, *J. Phys. Conf. Ser.* **898** (2017) 102006 [1704.05473].

[110] D. P. Kingma and J. Ba, *Adam: A method for stochastic optimization*, 2014.

[111] CMS collaboration, *Study of Z boson production in pPb collisions at* $\sqrt{s_{NN}} = 5.02$ *TeV*, *Phys. Lett. B* **759** (2016) 36 [1512.06461].

[112] ATLAS collaboration, *Measurement of prompt photon production in* $\sqrt{s_{\mathrm{NN}}} = 8.16$ *TeV p+Pb collisions with ATLAS*, *Phys. Lett. B* **796** (2019) 230 [1903.02209].

[113] H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, J. Pumplin et al., *New parton distributions for collider physics*, *Phys. Rev. D* **82** (2010) 074024 [1007.2241].

[114] K. Eskola, H. Paukkunen and C. Salgado, *EPS09 - Global NLO analysis of nuclear PDFs and their uncertainties*, 2009. 10.22323/1.080.0019.

[115] V. Bertone, S. Carrazza and J. Rojo, *Apfel: A pdf evolution library with qed corrections*, *Computer Physics Communications* **185** (2014) 1647–1668 [1310.1394].

[116] J. M. Campbell, J. Rojo, E. Slade and C. Williams, *Direct photon production and PDF fits reloaded*, *Eur. Phys. J. C* **78** (2018) 470 [1802.03021].

[117] S. Catani, M. Fontannaz, J. Guillet and E. Pilon, *Cross-section of isolated prompt photons in hadron hadron collisions*, *JHEP* **05** (2002) 028 [hep-ph/0204023].

[118] CMS collaboration, *Constraining gluon distributions in nuclei using dijets in proton-proton and proton-lead collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV*, *Phys. Rev. Lett.* **121** (2018) 062002 [1805.04736].